

## **MÉMOIRE**

*pour le*

**Diplôme d'Études Approfondies**

**« Méthodes de Recherche sur l'Environnement et la Santé »**



---

**Évaluation des performances du test de diagnostic  
cELISA utilisé pour le suivi de la fièvre catarrhale  
du mouton en France  
Conséquences sur la prévalence réelle de la maladie  
et sur la stratégie d'échantillonnage**

soutenu le 30 juin 2003

*par*

**Fabienne Biteau-Coroller**

**Année universitaire 2002-2003**

Laboratoire d'accueil : Programme « Santé Animale » ;

CIRAD-EMVT (Centre de coopération internationale en recherche  
agronomique pour le développement), Montpellier

Chef de programme : Emmanuel Camus

Maître de stage : François Roger

# TABLE DES MATIÈRES

<i>Remerciements</i> .....	2
<i>Résumé</i> .....	3
<i>Abréviations</i> .....	4
<i>Liste des tableaux, figures et annexes</i> .....	5
<i>Introduction</i> .....	8
 <b>1. <u>MATÉRIELS ET MÉTHODES</u></b> .....	<b>10</b>
<b>1.1. ÉVALUATION DES TESTS DE DIAGNOSTIC</b> .....	<b>10</b>
1.1.1. PRINCIPE GÉNÉRAL DE L'ÉTUDE .....	10
1.1.3.1. <i>Type d'étude</i> .....	10
1.1.3.2. <i>Statut de l'animal</i> .....	10
1.1.3.3. <i>Évolution de la virémie et du statut immunitaire après infection</i> --	11
1.1.2. DONNÉES .....	13
1.1.3. RECEIVER-OPERATING CHARACTERISTIC (ROC) ANALYSE .....	20
1.2.2.1. <i>Généralités</i> .....	20
1.2.2.2. <i>Démarche suivie</i> .....	23
1.2.2.3. <i>Logiciels utilisés</i> .....	23
<b>1.2. COMPARAISON DES DEUX TESTS CELISA</b> .....	<b>24</b>
1.2.1. PRINCIPE GÉNÉRAL DE L'ÉTUDE .....	24
1.2.2. DONNÉES .....	24
1.2.2.1. <i>Généralités</i> .....	24
1.2.2.2. <i>Sélection des sérums</i> .....	24
1.2.3. COEFFICIENT KAPPA .....	25
 <b>2. <u>RÉSULTATS</u></b> .....	<b>26</b>
<b>2.1. ÉVALUATION DU TEST DE DIAGNOSTIC CELISA I</b> .....	<b>26</b>
<b>2.2. ÉVALUATION DU TEST DE DIAGNOSTIC CELISA II</b> .....	<b>27</b>
<b>2.3. COMPARAISON DES DEUX TESTS</b> .....	<b>32</b>
2.3.1. SANS RÉAJUSTEMENT DU SEUIL .....	32
2.3.2. AVEC RÉAJUSTEMENT DU SEUIL .....	32
<b>2.4. EXEMPLES D'APPLICATIONS :CALCUL DE LA PRÉVALENCE APPARENTE CHEZ LES OVINS EN 2000</b> .....	<b>33</b>
 <b>3. <u>DISCUSSION</u></b> .....	<b>34</b>
<i>Références bibliographiques</i> .....	38
ANNEXES .....	41

## Remerciements

Mes remerciements vont à tous ceux et toutes celles qui ont contribué de près ou de loin à l'accomplissement de ce présent mémoire.

Je remercie sincèrement François Roger qui m'a suivi et orienté tout au long de ce stage ; je lui suis particulièrement reconnaissante pour son aide éclairée, ses précieux conseils et sa disponibilité.

Des remerciements appuyés vont tout naturellement à Colette Grillet qui a conduit tout le volet expérimental de l'étude. Merci pour son accueil au sein du laboratoire et ses indispensables renseignements techniques et pratiques.

Merci à Emmanuel Albina, responsable de l'équipe virologique, qui a appuyé notre étude.

J'adresse un grand merci à David Chavernac qui m'a été d'un grand recours lors de la phase préliminaire de collecte des données, pour son aide précieuse et ses conseils fructueux tout au long de l'élaboration de ce mémoire.

Je tiens également à remercier les personnes suivantes pour leur contribution, leur disponibilité et leur accueil :

Stefan Zientara et Corinne Sailleau grâce auxquels j'ai pu bénéficier des données traitées par l'AFSSA d'Alfort ;

Laure Alnot, correspondante à la DGA1 ;

Les différents acteurs sur le terrain :

- à la DSV de haute Corse : Catherine Bouvier, directrice, Alexandre Bouchot, adjoint santé animale et Joël Lefèbvre, technicien santé animale ;

- à la DSV de Corse du Sud : Jean-Michel Bergès, Directeur, responsable régional Corse, Jacques Parodi, adjoint santé animale, Guy Bousquet, technicien santé animale (antenne de Sartène) ainsi que les techniciens de la direction d'Ajaccio ;

- les techniciens des laboratoires départementaux de Corse pour l'énorme travail d'analyses qu'ils doivent mener dans des conditions parfois difficiles et qui est à la base du suivi épidémiologique de la fièvre catarrhale du mouton en Corse.

J'ai eu également le plaisir de m'initier aux systèmes d'information géographique grâce à aux enseignements délivrés par Stéphane de la Rocque. Merci encore pour sa sympathie, ses conseils éclairés et sa disponibilité.

Merci enfin à Mickaël pour sa relecture du mémoire, ses conseils pertinents et son soutien.



## Résumé

Depuis 1998, la fièvre catarrhale du mouton (FCM ou bluetongue) s'est progressivement propagée dans différents pays européens du pourtour méditerranéen tels la Grèce, la Bulgarie, l'Italie (Sardaigne, Sicile, Calabre) et l'Espagne (îles Baléares). En France, les premiers foyers cliniques ont été déclarés en Corse en octobre 2000. La présence du sérotype 2 du virus de la bluetongue était confirmée en 2001 avec un total de 335 foyers cliniques recensés. La lutte contre cette maladie réputée contagieuse chez les ovins s'articule autour de deux axes : des mesures de police et de prophylaxie sanitaires en Corse pour enrayer la maladie clinique et un programme de surveillance sérologique et entomologique en Corse et sur le continent français débuté en 2001. Ces deux volets s'inscrivent dans le programme d'éradication présenté par la France à la Commission Européenne. Les tests de diagnostic sérologique (tests cELISA) sont ainsi au cœur du dispositif de veille et de suivi de la circulation virale.

Les objectifs de l'étude étaient d'une part d'évaluer les performances (sensibilité et spécificité) des deux tests ELISA de compétition (cELISA I en 2000 et cELISA II en 2001) utilisés pour le diagnostic et le suivi de l'infection de la FCM en Corse en 2000 et 2001 et d'autre part de les comparer.

L'approche retenue pour conduire cette évaluation est l'analyse ROC [2, 4]. Elle s'est basée sur les données sérologiques et virologiques (détection de l'ARN viral par RT-PCR) collectées lors des deux campagnes de prophylaxie de 2000 et 2001. Le test de référence (*gold standard*) retenu pour définir le statut des animaux vis-à-vis de l'infection a été la RT-PCR. 406 sérums prélevés en 2000 (cELISA I) ont été également analysés de nouveau avec le test cELISA II. Un test kappa a été appliqué aux résultats afin de mettre en évidence le niveau de concordance entre les deux tests.

Les aires sous la courbe ROC obtenues sont de 0,84 (IC à 95 % : 0,73 – 0,95) pour le test cELISA utilisé en 2000 et de 0,78 (IC à 95 % : 0,68 – 0,89) pour celui utilisé en 2001. Les meilleures valeurs de sensibilité et de spécificité sont atteintes pour une valeur-seuil définie respectivement à 70 % et à 53 %. A la lumière de ces résultats, la prévalence sérologique observée en 2000 a été revue. Les valeurs de Kappa calculées respectivement pour un seuil à 50 % et à 70 %, pour le test cELISA I, sont de 0,0058 ( $\sigma^2 = 0,0053$ ) et de 0,1898 ( $\sigma^2 = 0,0350$ ). Cette concordance est qualifiée de mauvaise d'après Brennan & Hays (1992).

Les performances observées pour ces tests dans les conditions de sérodiagnostic sont discutées en particulier par rapport à l'utilisation de la RT-PCR comme test de référence. En effet, si ce test est reconnu excellent du point de vue de sa sensibilité et de sa spécificité, il ne détecte pas les anticorps mais l'ARN viral. Or, la virémie est précoce (2 à 3 jours) par rapport à la phase de production d'anticorps détectés par le test cELISA (minimum de 10 jours) [Koumbati *et al.*, 1999]. Le choix d'un test de référence sérologique telle la séroneutralisation apparaît ainsi essentiel pour compléter ces résultats préliminaires. La qualité des données intégrées à l'étude fait également l'objet d'une réflexion. Les seules données disponibles étaient en effet celles collectées lors de la gestion des épizooties corses. Le problème de la représentativité et de la qualité de ces données qui n'ont pas été produites dans le but de mener une telle étude est alors posé. Finalement, des données supplémentaires et notamment expérimentales sont nécessaires afin de préciser les qualités intrinsèques de ces tests particulièrement pour leur utilisation comme outil de vigilance dans les zones indemnes. La comparaison des deux tests s'est heurtée quant à elle à un fort déséquilibre dans la répartition des résultats en fonction du test. Il s'ensuit une difficulté à commenter les résultats du test kappa. L'échantillon sera complété afin de pouvoir renforcer la puissance de ce test.



## Abréviations

AHSV :	African Horse Sickness Virus (virus de la maladie équine virale africaine)
ARN :	Acide RiboNucléique
AUC :	Area Under the Curve (aire sous la courbe)
BTV :	Blue Tongue Virus (virus de la bluetongue)
cELISA :	competitive Enzyme-Linked ImmunoSorbent Assay
DEA :	Diplôme d'Études Approfondies
EHDV :	Epizootic Hemorrhagic Disease Virus (virus de la fièvre hémorragique du cerf)
ELD <sub>50</sub> :	Egg Lethal Dose 50 (Dose entraînant 50 % de mortalité des oeufs embryonnés)
FCM :	Fièvre catarrhale du mouton
FN :	Faux Négatifs
FP :	Faux positifs
IC :	Intervalle de Confiance
P :	Prévalence de la maladie
Pa :	Prévalence apparente de la maladie
Pr :	Prévalence réelle de la maladie
ROC :	Receiver Operating Characteristic
RT-PCR:	Reverse Transcriptase – Polymerase Chain Reaction (épreuve d'amplification en chaîne par polymérase)
Sd :	écart-type (standard deviation)
se :	sensibilité
sp :	spécificité
TG-ROC :	Two Graph – Receiver Operating Characteristic
VN :	Vrais Négatifs
VP :	Vrais Positifs
VPP :	Valeur Prédictive Positive
VPN :	Valeur Prédictive Négative

## Liste des tableaux, figures et annexes

### Figures :

Figure 1 :	Exemple d'évolution moyenne du titre viral dans le sang de moutons après inoculation du BTV, [Luedke, 1969].....	11
Figure 2 :	Évolution du taux d'anticorps mesuré par un test cELISA suite à une infection par le virus de la bluetongue (BTV), [Afshar <i>et al.</i> , 1987 ].....	12
Figure 3 :	Chronologie des événements virémique, clinique et sérologique de la fièvre catarrhale du mouton .....	13
Figure 4 :	Répartition communale du cheptel ovin en Corse en 2000 .....	14
Figure 5 :	Localisation des prélèvements sérologiques effectués en Corse en 2000 .....	14
Figure 6 :	Distribution des valeurs de pourcentage d'inhibition obtenues avec le test cELISA I chez les ovins prélevés en 2000 (n = 8 825) .....	15
Figure 7 :	Diagramme à moustaches représentant la dispersion des valeurs de pourcentage d'inhibition obtenues avec le test cELISA I chez les ovins en 2000.....	15
Figure 8 :	Distribution des valeurs de pourcentage d'inhibition de l'échantillon n1 correspond aux 64 sérums prélevés en 2000 et ayant à la fois un résultat sérologique (test cELISA I) et virologique (RT-PCR) .....	16
Figure 9 :	Diagramme à moustaches représentant la dispersion des valeurs de pourcentages d'inhibition dans l'échantillon n1 .....	16
Figure 10 :	Distribution des valeurs de pourcentage d'inhibition obtenues avec le test cELISA II chez les ovins prélevés en 2001 en Corse (n = 2607).....	17
Figure 11 :	Diagramme à moustaches représentant la dispersion des valeurs de pourcentages d'inhibition obtenues avec le test cELISA II chez les ovins en 2001 .....	17
Figure 12 :	Distribution des valeurs de pourcentage d'inhibition de l'échantillon n2 correspond aux 338 sérums prélevés en 2001 et ayant à la fois un résultat sérologique (test cELISA II) et virologique (RT-PCR) .....	18
Figure 13 :	Diagramme à moustaches représentant la dispersion des valeurs de pourcentages d'inhibition dans l'échantillon n2 .....	18
Figure 14 :	Distribution des valeurs de pourcentage d'inhibition obtenues pour les sérums de l'échantillon n2 provenant d'animaux non vaccinés (n = 104).....	19
Figure 15 :	Distribution des valeurs de pourcentage d'inhibition obtenues avec le test cELISA II sur des bovins prélevés en 2002 (n = 447) en zone indemne de FCM (continent français).....	19
Figure 16 :	Différentes allures de courbe ROC en fonction du pouvoir discriminant du test.....	20
Figure 17 :	Représentation graphique de l'effet de la variation du seuil sur la proportion des faux-positifs et des faux-négatifs.....	21
Figure 18 :	Exemple de courbe TG-ROC [Greiner, 1995 & 1996] .....	22
Figure 19 :	Formule de la prévalence réelle d'une maladie, [ Rogan & Gladen, 1978].....	22
Figure 20 :	Définitions des valeurs prédictives positives (VPP) et négatives (VPN) (d'après le théorème de Bayes) .....	22
Figure 21 :	Courbe ROC obtenue avec le test cELISA I ( n = 64) .....	26
Figure 22 :	Courbe ROC obtenue en retenant seulement les sérums provenant d'animaux non vaccinés (n <sub>2</sub> ' = 103, cELISA II) .....	27

Figure 23 :	Courbe ROC obtenue en conservant dans l'échantillon que les animaux non vaccinés dont le motif de prélèvement était la suspicion clinique ( $n_2'' = 85$ , cELISA II) .....	27
Figure 24 :	Courbe ROC obtenue en retenant que les sérums provenant d'ovins non vaccinés ( $n_2''' = 87$ , cELISA II) .....	28
Figure 25 :	Courbe ROC obtenue en retenant que les sérums provenant d'ovins non vaccinés dont le motif de prélèvement était une suspicion clinique de fièvre catarrhale du mouton ( $n_2'''' = 76$ , test cELISA II) .....	28
Figure 26 :	Courbe TG-ROC permettant de mettre en évidence le cut-off correspondant aux meilleures valeurs de sensibilité et de spécificité, le poids de ces deux paramètres étant considéré comme équivalent. ....	29
Figure 27 :	Courbe ROC obtenue en ajoutant les 447 résultats de bovins provenant d'une zone indemne de FCM (continent français, campagne 2002) au 103 sérums corses ( $n_2'''' = 550$ , cELISA II).....	30
Figure 28 :	Courbe ROC obtenue en ajoutant les 447 résultats de bovins provenant d'une zone indemne de FCM (continent français, campagne 2002) aux 87 sérums ovins corses ( $n = 534$ , cELISA II) .....	30
Figure 29 :	Courbe ROC obtenue en ajoutant les 3 755 résultats de bovins provenant d'une zone indemne de FCM (continent français, campagne 2001 et 2002) aux 103 sérums corses ( $n = 3858$ , cELISA II) .....	30
Figure 30 :	Boîte à moustaches représentant la dispersion des pourcentages d'inhibition (PI) , $n = 406$ , en fonction du test utilisé (cELISA I vs cELISA II) .....	32
Figure 31 :	Prévalence apparente chez les ovins en 2000 avec un seuil défini à 50 % pour le test cELISA I .....	33
Figure 32 :	Prévalence apparente chez les ovins en 2000 avec un seuil défini à 70 % pour le test cELISA I .....	33

### Tableaux :

Tableau I :	Contingence des animaux selon leur statut vis-à-vis de l'infection et leur résultat en cELISA – définition des paramètres.....	11
Tableau II :	Nombre de sérums collectés en 2000 pendant l'épizootie corse de bluetongue .....	15
Tableau III :	Distribution des sérums prélevés en 2000 selon le résultat sérologique (cELISA I) .....	24
Tableau IV :	Détail concernant les sérums de 2000 recontrôlés par le test cELISA II.....	25
Tableau V :	Synthèse de l'ensemble des résultats de l'analyse ROC en fonction des critères de sélection des sérums .....	31
Tableau VI :	Résultats obtenus avec les sérums prélevés en 2000 ( $n = 406$ ) selon le test cELISA utilisé, les deux tests ayant une valeur-seuil fixée à 50 %. ....	32
Tableau VII :	Résultats obtenus avec les sérums prélevés en 2000 ( $n = 406$ ) selon le test cELISA utilisé, le test cELISA I ayant une valeur-seuil fixée à 70 %. ....	33



## Annexes :

ANNEXE I :	Localisation des derniers foyers recensés de fièvre catarrhale du mouton dans le bassin méditerranéen (données 2002) – source OIE .....	42
ANNEXE II:	Manual of standards Diagnostic Test and Vaccines 2000, Principles of validation of diagnostic assays for infectious diseases.....	43
ANNEXE III :	Morphologie du virus de la fièvre catarrhale du mouton .....	52
ANNEXE IV :	Base de données relationnelles créée pour la gestion des données sérologiques et virologiques collectées dans le cadre de la surveillance de la bluetongue en Corse.....	53
ANNEXE V :	Détail des résultats de l'analyse ROC conduite en retenant dans l'échantillon les animaux non vaccinés dont le motif de prélèvement était la suspicion clinique .....	54
ANNEXE VI :	Manuel opératoire du test cELISA II .....	56

## Introduction

En octobre 2000, les premiers foyers de fièvre catarrhale du mouton<sup>1</sup> (Bluetongue) ont été recensés dans la région du Sartenais et dans la plaine de Porto Vecchio en Corse. Cette maladie est classée dans la liste A de l'OIE (Office International des Épizooties) c'est-à-dire qu'elle fait partie des « maladies transmissibles qui ont un grand pouvoir de diffusion et une gravité particulière, susceptible de s'étendre au-delà des frontières nationales, dont les conséquences socio-économiques ou sanitaires sont graves et dont l'incidence sur le commerce international des animaux et des produits d'origine animale est très importante ».

Un ensemble de mesures de police sanitaire visant à circonscrire et à stopper la diffusion du virus (interdiction d'exporter tous les animaux des espèces sensibles – ruminants – depuis la Corse, abattage des animaux malades dans les élevages reconnus infectés etc.) a été déployé. Afin de recouvrer un statut indemne de fièvre catarrhale du mouton, un programme d'éradication de la maladie et de l'infection (« programme national d'éradication » soumis annuellement à la Commission Européenne) a été dès 2002 proposé et appliqué par les services vétérinaires français<sup>2</sup>. Ce programme est basé sur un système de surveillance sérologique et entomologique qui doit permettre, en l'absence de foyers cliniques, de démontrer la disparition de toute circulation virale.

Le dernier foyer recensé de fièvre catarrhale du mouton (FCM) en Corse a été confirmé le 8 novembre 2001. Cependant, en 2003, la Corse est toujours déclarée infectée par la fièvre catarrhale du mouton puisque la circulation virale de la maladie, en l'absence de foyers cliniques, a été démontré pour l'année 2002. Or, au vu de la réglementation internationale (Code Zoosanitaire International), un pays (ou une zone) est déclaré infecté par le virus si « la présence du virus a été signalée au cours des deux dernières années ». En revanche, la France continentale reste une zone indemne de la maladie car elle n'est pas adjacente à une zone infectée (le bassin méditerranéen français est situé à plus de 150 km de la Corse).

La fièvre catarrhale du mouton est une maladie virale, virulente, non contagieuse, saisonnière, transmise par des insectes hématophages. Elle touche les ruminants domestiques et sauvages. Les infections cliniques se manifestent essentiellement chez les ovins. Elle se caractérise par une forte fièvre, un jetage et un ptyalisme abondant, un œdème et des ulcérations de la muqueuse buccale. Une cyanose de la langue est parfois observée. Dans les tableaux les plus sévères, des raideurs musculaires ou un torticolis ainsi que des boiteries associées à une congestion des extrémités podales sont régulièrement décrits. Le taux de mortalité varie alors de 2 à 30 % [Verwoerd & Erasmus, 1994]. Après une longue convalescence, les survivants demeurent des non valeurs économiques (retard de croissance, mauvaise qualité de la laine, déclassement des carcasses).

---

<sup>1</sup> La dénomination de « fièvre catarrhale du mouton » (FCM) est le nom français officiel de cette maladie, nom qui figure dans la liste A du code zoosanitaire international en langue française. Le terme privilégié pour cette maladie au niveau international est « bluetongue », nom anglais qui fait référence à la cyanose de la langue qui est parfois observée dans le tableau clinique des ovins.

Le terme de fièvre catarrhale ovine a également fait son apparition sous l'impulsion des gestionnaires de l'épizootie afin d'éviter une confusion avec la fièvre catarrhale maligne auprès des éleveurs. Cependant, ce terme créé *ad hoc*, s'il peut être rencontré dans les rapports émanant de la Direction Générale de l'Alimentation, n'a pas de fondements scientifiques.

<sup>2</sup> En France, les services vétérinaires sont sous la tutelle du Ministère de l'Alimentation, de l'Agriculture, de la Pêche et des Affaires Rurales (MAAPAR) et s'organisent autour d'une Direction générale de l'Alimentation (DGAI). Dans chaque département se situe des services déconcentrés, les Directions Départementales des Services Vétérinaires (DDSV), chargés notamment de la mise en place des mesures de police et de prophylaxie sanitaires en cas de maladies réputées contagieuses.



L'extrême variabilité des manifestations cliniques s'explique notamment par une sensibilité différente des races ovines. Les conditions d'élevage, l'environnement jouent également un rôle déterminant [Verwoerd & Erasmus, 1994 ; Lefèvre, 2003]. Chez les autres espèces de ruminants (bovins, caprins, etc.), l'infection demeure généralement asymptomatique. Ils multiplient le virus mais cette virémie est le plus souvent sans répercussions sur l'état général de l'animal. Très rarement on peut observer un tableau clinique comparable à celui des ovins sur des caprins et exceptionnellement chez les bovins.

Cette maladie est due à un virus de la famille des *Reoviridae*, genre *Orbivirus*, comprenant 24 sérotypes. Il est transmis par des diptères du genre *Culicoïdes*, famille des *Ceratopogonidae* [Meiswinkel, 1994]. Le vecteur principal en Afrique (zone d'enzootie) ou dans certains pays du bassin méditerranéen touchés épisodiquement par la maladie (épizootie de 1956-1960 au Portugal et en Espagne) est *Culicoïdes imicola*. *Culicoïdes obsoletus* est parfois suspecté dans des épizooties touchant le bassin méditerranéen (Chypre en 1977). Cette espèce est ainsi considérée comme probablement vectrice. Le virus BTV a notamment déjà été isolé à partir de cette espèce mais la transmission n'a jamais pu être démontrée [Mellor & Pitzolis, 1979 ; Lefèvre, 2003]. Comme la plupart des insectes, la dynamique de population de ces vecteurs est très largement influencée par les conditions environnementales et notamment climatiques. Les paramètres tels que l'humidité, la température, la végétation, etc. sont autant de facteurs importants influençant la répartition de ces vecteurs et donc celle du virus. A l'heure actuelle, beaucoup d'inconnues persistent quant aux caractéristiques morpho-écologiques des gîtes larvaires de *Culicoïdes imicola*.

L'aire de répartition de ce virus était, jusqu'à une période récente, délimitée au nord par le 40 / 50<sup>ème</sup> parallèle et au sud par le 35<sup>ème</sup> parallèle. Toutefois, à partir de 1998, il a commencé à se propager durablement dans le bassin méditerranéen atteignant la Grèce en 1998 et en 1999, la Bulgarie en 1999, l'Italie (Sardaigne, Sicile, Calabre) et l'Espagne (îles Baléares) en 2000 (Annexe I). Ainsi, en France continentale, les régions Languedoc-Roussillon et Provence-Alpes-Côte d'Azur se situant à proximité de régions infectées apparaissent comme des zones à risque d'émergence du virus d'autant plus qu'elles abritent des écosystèmes qui seraient favorables au vecteur *Culicoïdes imicola* d'après les prédictions obtenues à partir d'un modèle mathématique basé sur les données corses et des variables environnementales [Roger et al., 2002]. Cette approche par modélisation permet de proposer une extension des zones de surveillance et incite à promouvoir une coopération méditerranéenne dans ce domaine.

En tant que maladie de la liste A du code zoosanitaire international de l'OIE, elle fait l'objet de recommandations spécifiques pour la sécurisation des échanges entre les pays en fonction de leur statut vis-à-vis de cette maladie. Les tests de diagnostic préconisés par l'OIE sont la PCR, l'isolement viral, l'ELISA de compétition et l'immunodiffusion sur gélose [OIE, 2000]. Le test de séroneutralisation est également un test de référence matière mais il détecte les anticorps spécifiques du type et non du groupe des BTV. Il est ainsi plutôt conseillé si le sérotype du virus de la bluetongue recherchée est connu.

Dans le volet sérologique du dispositif français de veille (continent) et de suivi de la circulation virale (Corse), le test de diagnostic utilisé (cELISA de Pirbright puis cELISA de VMRD Inc, USA) occupe une place centrale. L'objectif principal du stage de DEA était d'évaluer les caractéristiques intrinsèques (sensibilité, spécificité) du test sérologique actuellement utilisé et d'évaluer la possibilité de définir une valeur-seuil optimale pour ce test. Les applications possibles sont multiples : ces données permettront en particulier d'affiner l'interprétation des résultats sérologiques en fonction de la situation épidémiologique qui n'est pas la même sur le continent ou en Corse, d'estimer la prévalence réelle de la maladie en Corse et de redéfinir la taille des échantillons du programme de surveillance. Compte tenu du fait que deux tests sérologiques ont été successivement utilisés, une partie du stage a également été consacrée à la comparaison de ces tests.



# 1. Matériels et Méthodes

## 1.1. Évaluation des tests de diagnostic

### 1.1.1. Principe général de l'étude

#### 1.1.3.1. Type d'étude

La validation des tests de diagnostic est essentielle dans leur processus de mise au point. Compte tenu des carences générales dans ce domaine, les principes à respecter lors de la conduite d'une étude visant la validation d'un test de diagnostic ont été énoncés dans le manuel des standards des tests de diagnostic et des vaccins de l'OIE [1] dès 1996 (Cf. annexe II). Ce manuel de référence en matière d'échanges internationaux d'animaux et produits d'animaux précise que l'étude se découpe en 5 étapes comprenant une étape de caractérisation des performances du test (sensibilité, spécificité) et de détermination de la valeur-seuil (étape 3). En raison du nombre très important de variables influençant les performances de sérodiagnostics des tests, il est de plus recommandé (étape 5) de revalider le test sur une population animale avec des particularités différentes.

Greiner et Gardner ont présenté en 2000 une synthèse des différentes approches possibles pour valider un test de diagnostic. Ils mettent en avant la nécessité, avant de lancer une étude pour la validation d'un test, de bien définir quels sont les objectifs de l'étude et du test, de sélectionner au préalable les méthodes d'analyses appropriées et enfin d'identifier les contraintes et les biais éventuels.

L'objet de l'étude est de déterminer la valeur intrinsèque du test ELISA de compétition utilisé dans le diagnostic de la fièvre catarrhale du mouton en France lors des épizooties de 2000 et 2001. Pour un test quantitatif, l'estimation des paramètres de sensibilité et de spécificité dépend de la valeur-seuil retenue. C'est pourquoi il peut être interprété comme un couple de paramètres à l'aide d'une analyse ROC [Greiner et Gardner, 2000].

En l'absence de test de référence parfait (« gold standard<sup>3</sup> »), la RT-PCR est utilisée comme test de référence pour définir le statut de l'animal vis-à-vis de l'infection. Cette technique permet de détecter l'ARN du virus de la FCM pendant au moins 30 jours après isolement viral [MacLachlan *et al.*, 1994] et jusqu'à 90 jours. Ce test est recommandé par l'OIE dans le cadre de la maîtrise des risques sanitaires liés aux échanges internationaux.

En fonction du statut de l'animal ainsi défini, les valeurs de pourcentage d'inhibition calculées du test cELISA seront analysées au moyen d'une analyse ROC (Cf. 1.1.3.).

#### 1.1.3.2. Statut de l'animal

L'infection d'un ruminant par le virus de la bluetongue ne signifie forcément pas que l'animal va développer la maladie. En effet, dans de nombreux cas, l'infection reste asymptomatique. L'espèce ovine est l'espèce qui est la plus sensible au virus et qui exprime le plus fréquemment des signes cliniques (forte température, ptyalisme, oedème de la face etc.) [Coetzer *et al.*, 1994]. Cependant, il existe une forte variation intra-espèce. Ainsi en Corse, pendant les épizooties de 2000 et 2001, la morbidité moyenne dans les élevages infectés fut comprise entre 20 % et 25 %, avec des variations de 0,3 % à 100 % et des tableaux cliniques plus ou moins sévères.

Le statut de l'animal vis-à-vis de l'infection ne peut ainsi pas être définie seulement par la présence ou non de signes cliniques de la maladie d'autant plus que ces derniers sont très variables et non pathognomoniques.

---

<sup>3</sup> Le « gold standard » que l'on peut traduire en français par « test de référence parfait » est un test dont les paramètres de sensibilité et de spécificité sont de 100 %. Il ne génère aucuns faux-positifs et aucuns faux-négatifs. En pratique, il existe très peu de vrai « gold standard ». Le test de référence est souvent le test qui s'en rapproche le plus.

Les anticorps détectés par le test cELISA peuvent être deux origines : soit ils sont liés à une infection naturelle, soit ils sont le fait d'une vaccination. Début 2001, une campagne de vaccination a été conduite en Corse. Compte tenu de la qualité des données concernant cette vaccination, seuls les animaux non vaccinés (essentiellement des agnelles de l'année) seront retenus.

Au vu du test de référence retenu, le statut de l'animal est défini de la manière suivante :

- Un animal est infecté si son prélèvement de sang et/ou d'organes donne un résultat positif en RT-PCR.
- Un animal non infecté donne un résultat négatif en RT-PCR.

**Tableau I : Contingence des animaux selon leur statut vis-à-vis de l'infection et leur résultat en cELISA – définition des paramètres**

Résultat du test cELISA	Statut vis-à-vis de la bluetongue		total
	RT-PCR positive (infecté)	RT-PCR négative (non infecté)	
Positif	a	b	a + b
Négatif	c	d	c + d
Total	a + c	b + d	a + b + c + d

a = animaux infectés donnant un résultat positif en cELISA (vrais positifs)

b = animaux non infectés donnant un résultat positif (faux positifs)

c = animaux infectés donnant un résultat négatif en cELISA (faux négatifs)

d = animaux non infectés donnant un résultat négatif en cELISA (vrais négatifs)

**Sensibilité** (se) =  $a / (a + c)$  avec

**Écart-type** (sd) =  $se * (100 - se) / n_{(a+c)}$

**Prévalence apparente** =  $(a + b) / (a + b + c + d)$

**spécificité** (sp) =  $d / (b + d)$  avec

**sd** =  $sp * (100 - sp) / n_{(b+d)}$

**prévalence réelle** =  $(a + c) / (a + b + c + d)$

### 1.1.3.3. Évolution de la virémie et du statut immunitaire après infection

La durée de la virémie est difficile à estimer car elle dépend de plusieurs facteurs (sérotypes, souches, individus, mode détection etc.). Le virus peut être isolé dans le sang d'un animal infecté à partir de 2 à 6 jours après l'infection [Bonneau *et al.*, 2002 ; Ghalib *et al.*, 1985 ; Hamblin *et al.*, 1998 ; Koumbati *et al.*, 1999 ; Luedke, 1969] avec un pic autour du 7<sup>ème</sup> jour post-infection [Ghalib *et al.*, 1985]. Chez les moutons, la virémie peut persister de 27 à 54 jours post-infection [Koumbati *et al.*, 1999]. Lefèvre (2003) considère qu'il est « raisonnable » d'admettre que la virémie chez les moutons est de 8 à 15 jours. L'ARN viral peut toutefois être détecté par RT-PCR pendant des périodes plus longues, supérieures à 100 jours [Bonneau *et al.*, 2002 ; Katz *et al.*, 1993 & 1994], même si cette positivité en RT-PCR n'a pas toujours de relation avec l'infectiosité de la particule virale [Katz *et al.*, 1994]. En effet, l'infection du vecteur n'est possible qu'au dessus d'un certain titre viral. Là encore, selon les études, le titre viral maximum varie sensiblement (exemple figure 1).

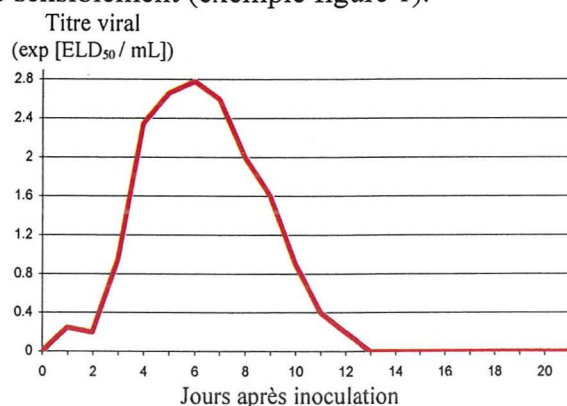


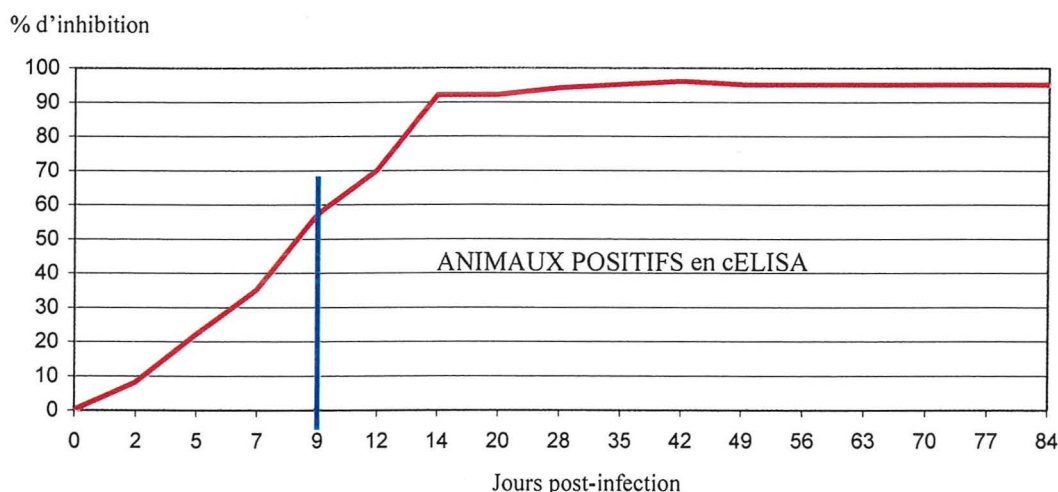
Figure 1 : Exemple d'évolution moyenne du titre viral dans le sang de moutons après inoculation du BTV, [Luedke, 1969]



Les moutons développent une protection immunitaire détectable par un test ELISA de compétition dès 6 à 10 jours post infection (figure 2) [Afshar *et al.*, 1987 ; Afshar *et al.*, 1989 ; Ghalib *et al.*, 1985 ; Hamblin *et al.*, 1998 ; Koumbati *et al.*, Uren *et al.*, 1982]. Le pic est atteint vers 4 semaines après l'infection [Ghalib *et al.*, 1985] et persiste au moins 6 mois [Lefèvre, 2003]. Les anticorps alors détectés sont des anticorps spécifiques du groupe BTV et sont dirigés contre la protéine structurale VP7. Cette protéine apparaît non seulement responsable de la l'antigénicité de groupe du virus de la bluetongue mais elle semble également contenir plusieurs déterminants antigéniques dont certains sont présents chez tous les sérotypes et d'autres sont seulement trouvés chez certains sérotypes [Campbell *et al.*, 1985]. Les antigènes communs entre le virus BTV et le virus de la fièvre hémorragique du cerf (EHSV) seraient portés par cette protéine VP7 et la protéine VP3 [Lefèvre, 2003] (Cf. annexe III).

Avant la mise au point des tests ELISA indirecte et ELISA de compétition, l'immunodiffusion en gélose (IDG) était le test sérologique de choix pour la détection des anticorps de groupe. Cependant, du fait notamment du manque de sensibilité et de l'existence de réactions croisées avec d'autres orbivirus tel le virus de la fièvre hémorragique du cerf, les techniques ELISA lui sont actuellement préférées.

La séroneutralisation reste le test sérologique de référence au niveau international. Elle détecte les anticorps neutralisants spécifiques de type. Ces anticorps neutralisent l'épitope de la protéine VP2 impliquée dans l'infectiosité de la particule virale. Des réactions croisées entre certains sérotypes existent.



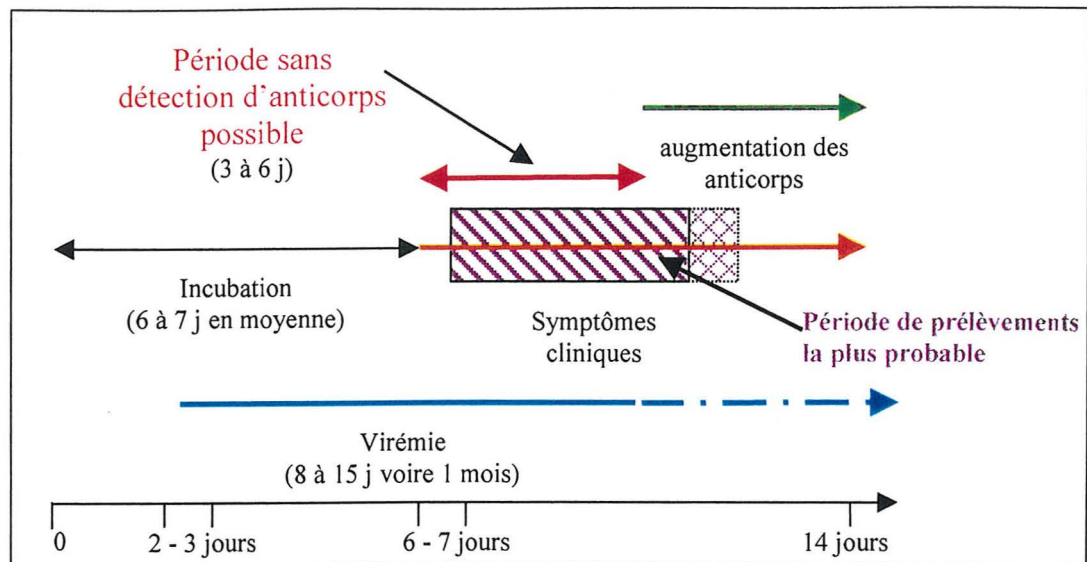
**Figure 2 : Évolution du taux d'anticorps mesuré par un test cELISA suite à une infection par le virus de la bluetongue (BTV), [Afshar *et al.*, 1987 ]**

Ces données sur l'évolution de la virémie et du taux d'anticorps dans le sang d'un mouton infecté permettent d'expliquer certains résultats faussement positifs ou négatifs.

Ainsi, des faux négatifs peuvent s'analyser de la manière suivante :

- l'infection par le virus BTV est récente (figure 3). L'animal n'a eu le temps de développer une protection immunitaire suffisante pour être détectée par un test sérologique (période de 7-10 jours post-infection).
- le prélèvement peut être de mauvaise qualité. Des problèmes dans l'interprétation des sérums fortement hémolysés ont ainsi été notés à plusieurs reprises au Cirad-emvt.
- La RT-PCR est très sensible aux contaminations de laboratoire qui peuvent conduire à des faux-positifs. Le défaut de sensibilité n'est dans ce cas pas imputable au test sérologique





**Figure 3 : Chronologie des événements virémique, clinique et sérologique de la fièvre catarrhale du mouton**

Les faux-positifs peuvent s'interpréter des façons suivantes :

- des réactions croisées avec d'autres virus ou avec d'autres facteurs non spécifiques sont à envisager [Afshar *et al.*, 1989]. Cependant, s'il est admis que l'immunodiffusion en gélose entraîne des réactions croisées avec le virus de la fièvre hémorragique du cerf, le test cELISA est connu pour ne pas croiser avec les anticorps de cet autre orbivirus (Afshar *et al.*, 1989). D'autres groupes de d'orbivirus sont proches antigéniquement comme le groupe des « African Horse Sickness » (ex : AHSV) ou le groupe des virus « Palyam ».
- l'animal a été en contact avec le virus lors de la première épidémie. Dans ce cas, il est considéré comme un faux-positif au vu de la définition du cas positif mais cela ne correspond pas à un défaut de spécificité du test. Ce cas de figure est notamment probable pour l'espèce caprine. Par contre, il ne peut être invoqué pour les agnelles.
- Dans le cas des agnelles nées de brebis vaccinées, la persistance d'anticorps maternels peut conduire à l'obtention de résultats positifs [Gaydos *et al.*, 2002].

### 1.1.2. Données

Un premier travail de récupération des données sérologiques et virologiques a été nécessaire en vue de leur exploitation. Une nouvelle base de données a été créée sous Access® afin de rassembler les différentes sources de données (CIRAD-EMVT, AFSSA-Alfort) et de permettre de faire des requêtes spécifiques en fonction des besoins. Au niveau du CIRAD-EMVT, les résultats sont saisis dans une base de données appelée Cora®.

En raison des informations disponibles et des modes de saisie particuliers à chacun des deux laboratoires impliqués, la mise en correspondance des résultats sérologiques et virologiques d'un même animal n'a pas été simple. Toutes les feuilles de commémoratifs archivées ont ainsi dû être vérifiées et les informations non saisies ou saisies sous un format différent entre les deux laboratoires ont été complétées et/ou corrigées.

Les données utilisées dans le cadre de cette étude proviennent des enquêtes épidémiologiques menées en Corse lors des deux épizooties de 2000 et 2001. Il s'agit ainsi dans la majorité des cas d'animaux cliniquement suspects. On étudie ainsi le test dans le cadre du sérodiagnostic d'une infection récente. Toutefois, des sérums collectés pendant les campagnes de prophylaxie bovine et caprine de l'hiver 2000/2001 ont également été analysés.

Pour l'année 2000, un premier test ELISA a été utilisé, cELISA I. 13 611 résultats sérologiques dont 8 825 provenant d'ovins (tableau II) sont ainsi disponibles. La première campagne de vaccination ayant débuté en janvier 2001, tous les sérums proviennent d'animaux non vaccinés.

La figure 4 montre la répartition du cheptel ovin en Corse. Les zones les plus prélevées en 2000 (figure 5) sont l'extrême sud de la Corse et quelques communes dispersées sur l'ensemble de l'île. La plaine orientale où l'élevage ovin est important n'a pas fait l'objet de beaucoup de prélèvements. Les zones d'analyses sont ainsi localisées sur les communes où des foyers cliniques de FCM se sont déclarés.

La figure 6 permet de visualiser la distribution des résultats sérologiques exprimés en pourcentage d'inhibition. La moyenne de cet échantillon est de 43,5 % avec un écart-type de 15,7 %. L'intervalle de confiance associé, pour  $\alpha = 0,05$ , est compris entre 43,5 % et 44,1 %. 25 % des sérums (figure 4) ont donné un résultat compris dans la zone douteuse du test (entre 45 et 55 % d'inhibition). Cette zone douteuse a été définie par le laboratoire national de référence (Cirad-emvt) *a posteriori* au regard de la distribution des premiers résultats. Dans le manuel de l'OIE [OIE, 2000], une zone douteuse comprise entre 40 et 50 % d'inhibition est également conseillée. Si le recours à la définition d'une zone douteuse permet augmenter le pouvoir discriminant d'un test, l'interprétation des résultats douteux n'en demeure pas moins délicate. Des investigations complémentaires sont alors nécessaires pour lever l'incertitude qui persiste sur ces résultats.

En début d'épizootie de FCM, la distribution des résultats sérologiques (figure 6) suit une loi normale. On ne distingue en effet pas deux populations distinctes d'animaux, ceux ayant développé une protection immunitaire suite au contact avec l'agent pathogène et ceux qui n'ont pas rencontré le virus. De plus, peu d'animaux obtiennent des pourcentages d'inhibition supérieurs à 75 %. Ces observations amènent à émettre l'hypothèse qu'au moment du prélèvement, ces animaux étaient plutôt dans une phase de séroconversion que dans une période où l'immunité humorale avait atteint son maximum.

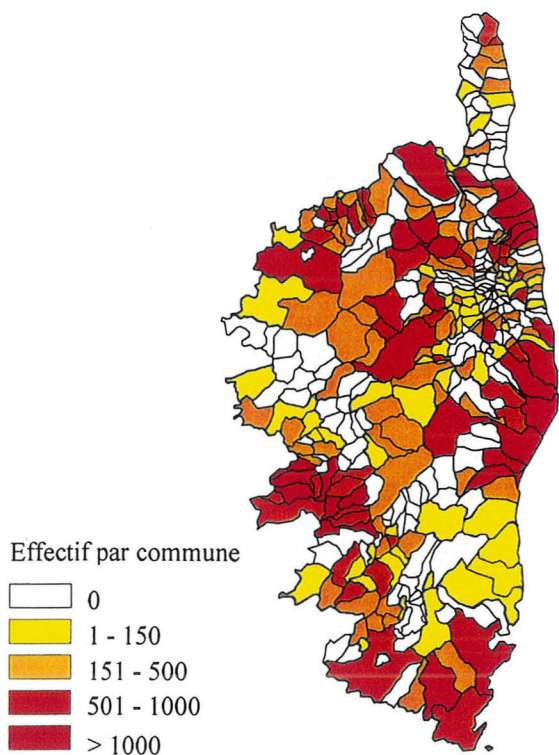


Figure 4 : Répartition communale du cheptel ovin en Corse en 2000

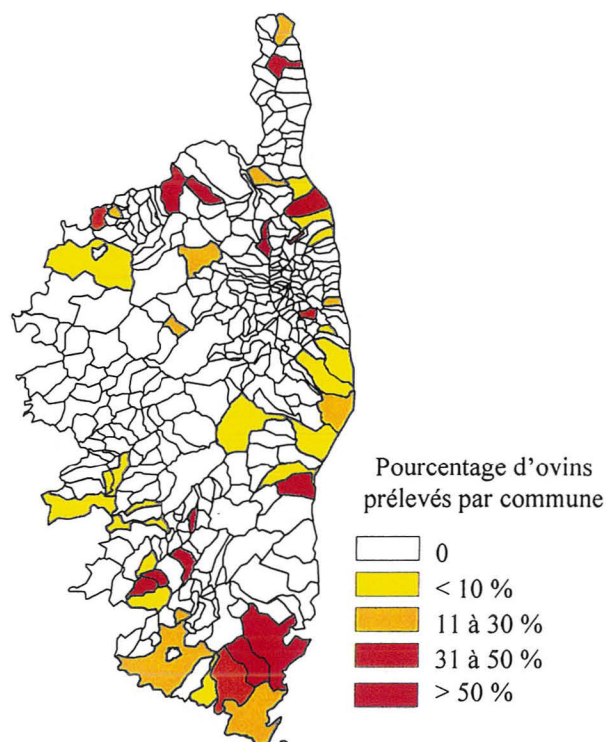
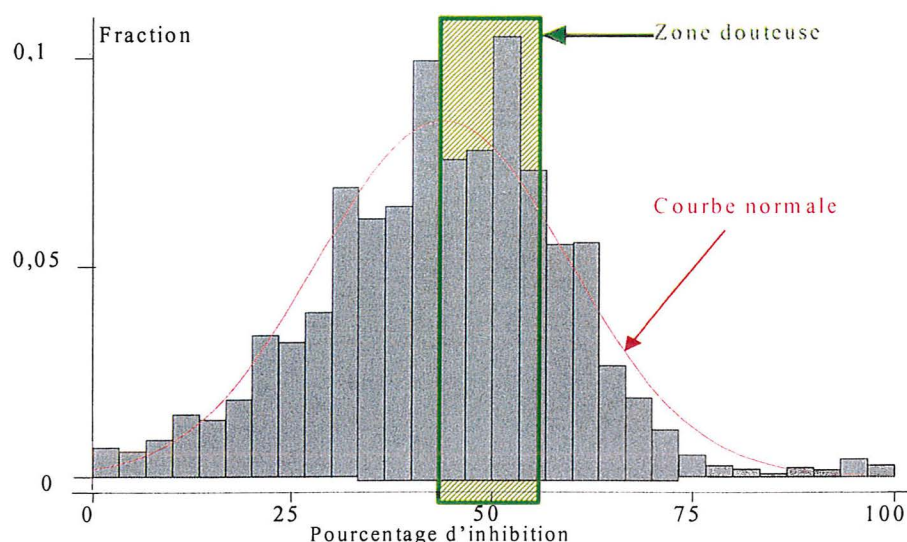


Figure 5 : Localisation des prélèvements sérologiques effectués en Corse en 2000

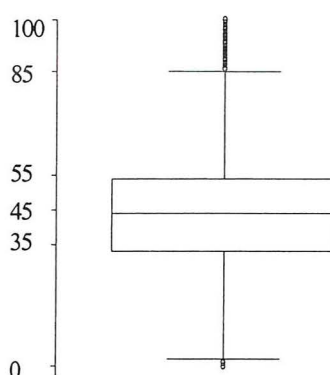


**Tableau II : Nombre de sérums collectés en 2000 pendant l'épizootie corse de bluetongue**

Espèce	Nombre de sérums	Pourcentage
BOVINS	4 637	34,1 %
CAPRINS	149	1,1 %
OVINS	8 825	64,8 %
Total	13 611	100,0 %



**Figure 6 Distribution des valeurs de pourcentage d'inhibition obtenues avec le test cELISA I chez les ovins prélevés en 2000 (n = 8 825)**

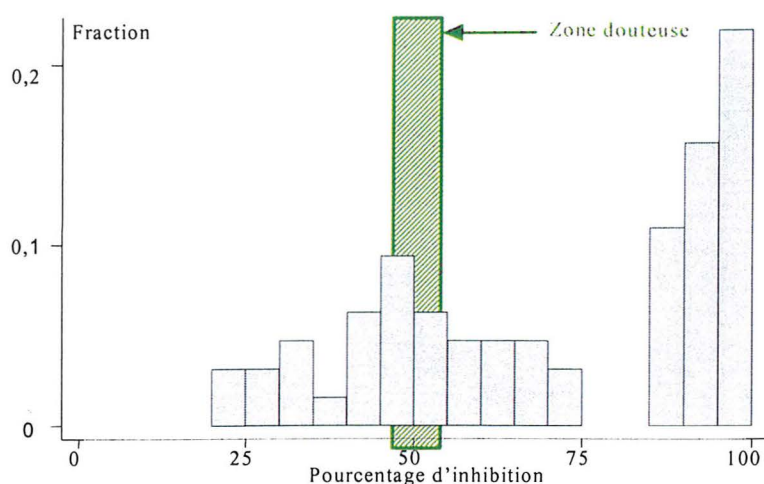


**Figure 7 : Diagramme à moustaches représentant la dispersion des valeurs de pourcentage d'inhibition obtenues avec le test cELISA I chez les ovins en 2000**

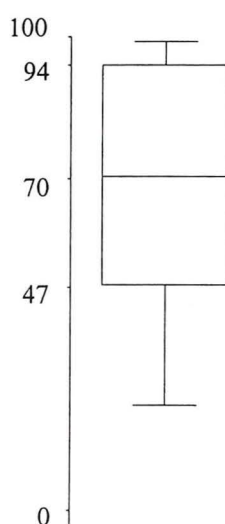
On dispose également de 193 résultats virologiques (RT-PCR). Parmi ces animaux testés, 64 ont fait à la fois l'objet d'une analyse sérologique et virologique. Ces sérums constituent l'échantillon  $n_1$  (figures 8 & 9). C'est à partir de cet échantillon qu'une analyse ROC a été réalisée (Cf. 1.1.3).

L'échantillon  $n_1$  se distribue différemment de l'échantillon total. La moyenne observée est de 69,9 % avec un intervalle de confiance compris entre 63,9 % et 75,9 %. Deux populations semblent se dessiner, une première dont les valeurs de pourcentage d'inhibition sont centrées autour de 50 % d'inhibition et une seconde pour laquelle les animaux obtiennent des résultats supérieurs à 80 % d'inhibition.





**Figure 8 :** Distribution des valeurs de pourcentage d'inhibition de l'échantillon  $n_1$  correspond aux 64 sérums prélevés en 2000 et ayant à la fois un résultat sérologique (test cELISA I) et virologique (RT-PCR)

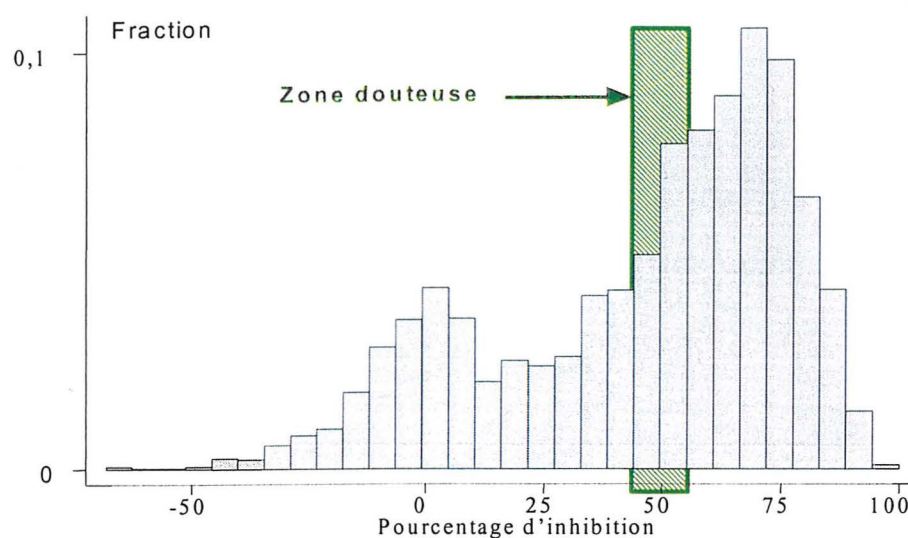


**Figure 9 :** Diagramme à moustaches représentant la dispersion des valeurs de pourcentages d'inhibition dans l'échantillon  $n_1$

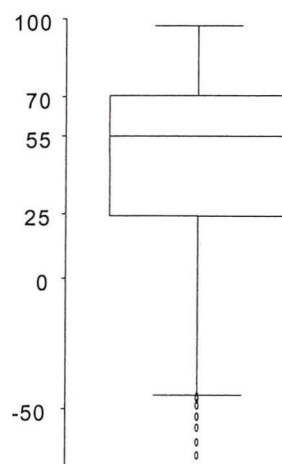
Pour l'année 2001, un autre test cELISA a été utilisé, noté cELISA II. Au total, 2 607 ovins ont fait l'objet d'une analyse sérologique, 736 d'une analyse virologique (RT-PCR) et 338 d'analyses virologique et sérologique.

La figure 10 représente la distribution des valeurs de pourcentage d'inhibition obtenues avec le test cELISA II sur les sérums prélevés en 2001 lors de suspicions cliniques de fièvre catarrhale du mouton ou lors du suivi d'élevages déclarés infectés. La moyenne dans cet échantillon est de 43,4 % avec un écart-type de 31,0 % (IC à 95 % : [42,2 % - 44,6 %]). La figure 10 précise la dispersion au sein de cet échantillon avec une médiane à 54,8 % et des percentiles 25 et 75 respectivement à 24,1 et 70,3 %.

La distribution en 2001 des résultats collectés n'est plus d'allure normale mais bi-modale. Pourtant, les conditions dans lesquelles ont été obtenus les prélèvements est sensiblement la même pour ces deux épizooties. Cependant, une partie de la population ovine corse n'est plus naïve vis-à-vis de la FCM. De plus, une campagne de vaccination a été menée jusque la fin du mois d'avril 2001. Ainsi, nous ne sommes plus dans le schéma d'un début d'épizootie comme en 2000. Une part des ovins a eu un contact assez ancien avec le virus ou la souche vaccinale pour manifester une couverture immunitaire plus importante.



**Figure 10 : Distribution des valeurs de pourcentage d'inhibition obtenues avec le test cELISA II chez les ovins prélevés en 2001 en Corse (n = 2607)**



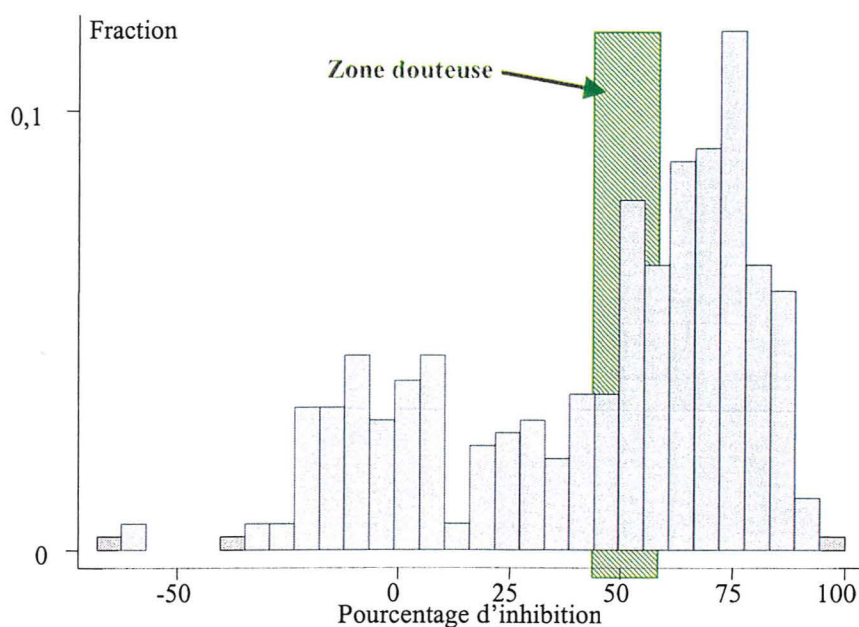
**Figure 11 : Diagramme à moustaches représentant la dispersion des valeurs de pourcentages d'inhibition obtenues avec le test cELISA II chez les ovins en 2001**

338 sérums disposent à la fois d'un résultat sérologique et virologique. Cet échantillon  $n_2$  est décrit par les figures 12 & 13.

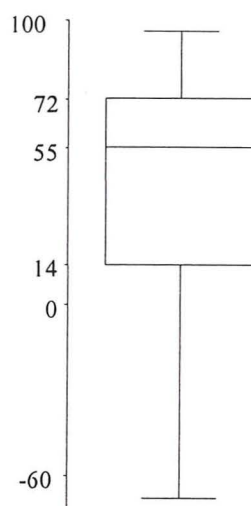
Les valeurs de pourcentage d'inhibition de l'échantillon  $n_2$  se distribuent également de manière bimodale (figure 12). Deux sous-populations semblent ainsi se dessiner, d'une part des animaux ayant développé une protection immunitaire contre le virus de la FCM et d'autre part des animaux non protégés. L'origine de cette protection est soit liée à la vaccination soit à une immunisation suite à une infection naturelle. En effet, dans ce lot de sérums, 103 proviennent d'animaux non vaccinés, 93 d'animaux vaccinés une fois et 142 d'animaux dont le statut vaccinal est inconnu.

La moyenne est de 43,9 % avec un écart-type de 35,4 % (IC à 95 % : [37,1 % - 50,7 %]). La médiane se situe à 55,2 % et 50 % des sérums sont compris entre 13,9 % et 72,4 % (figure 13).





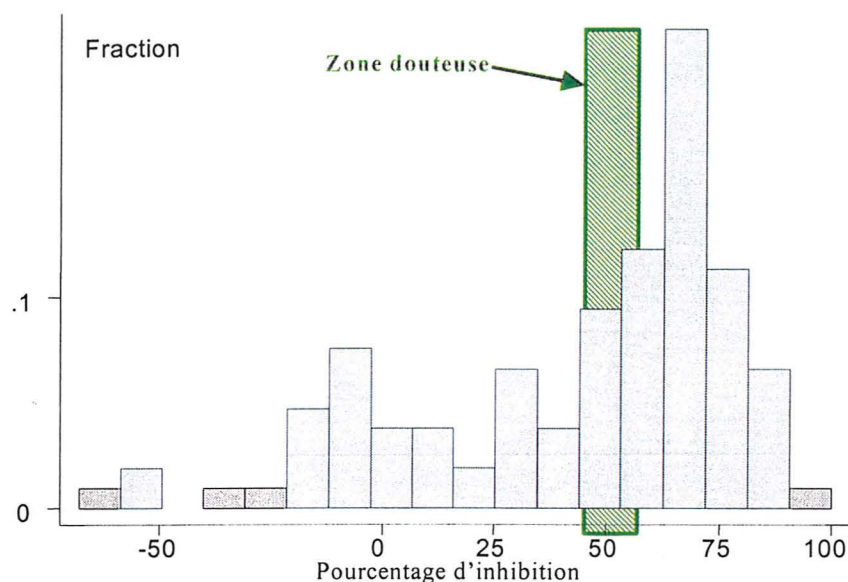
**Figure 12 : Distribution des valeurs de pourcentage d'inhibition de l'échantillon  $n_2$  correspond aux 338 sérums, prélevés en 2001, pour lesquels on dispose à la fois d'un résultat sérologique (test cELISA II) et virologique (RT-PCR)**



**Figure 13 : Diagramme à moustaches représentant la dispersion des valeurs de pourcentages d'inhibition dans l'échantillon  $n_2$**

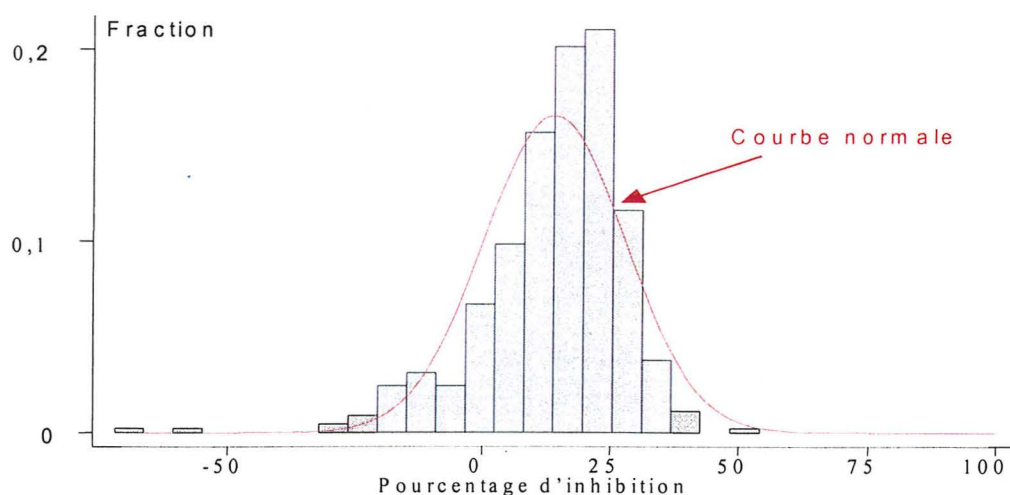
De cet échantillon, on conserve seulement les sérums provenant d'animaux non vaccinés soit 103 sérums (2 dromadaires<sup>4</sup>, 2 bovins, 12 caprins et 87 ovins). La distribution des valeurs de pourcentage d'inhibition (figure 14) conserve une allure bi-modale (2 sous-populations). La sous-population qui apparaît sérologiquement naïve vis-à-vis du virus obtient des pourcentages d'inhibition autour de 0 %. L'autre sous-population a des valeurs centrées autour de 70 % d'inhibition. Seulement 7 animaux obtiennent un résultat supérieur à 80 % d'inhibition.

<sup>4</sup> Le dromadaire est fréquemment trouvé séropositif vis-à-vis de la bluetongue en Afrique [Abu Elzein, 1985 ; Simpson, 1979] et pourrait jouer un rôle de réservoir.



**Figure 14 : Distribution des valeurs de pourcentage d'inhibition obtenues pour les sérums de l'échantillon  $n_2$  provenant d'animaux non vaccinés ( $n = 103$ )**

Enfin, on dispose de résultats sérologiques obtenus à partir de sérums prélevés sur des animaux situés en zone indemne de la maladie c'est-à-dire localisés en France sur le pourtour du bassin méditerranéen (programme de surveillance sur le continent). Le statut de ces animaux vis-à-vis de la bluetongue est ainsi connu. Cet échantillon est décrit par la figure 15. Les valeurs de pourcentage d'inhibition se distribuent autour de 14,1% (IC à 95 % : [12,8 % - 15,4 %] et écart-type = 13,8 %). Le statut de ces animaux par rapport à la maladie étant connu, ces données pourront être ajoutées, dans un second temps, à l'échantillon  $n_2$ .



**Figure 15 : Distribution des valeurs de pourcentage d'inhibition obtenues avec le test cELISA II sur des bovins prélevés en 2002 ( $n = 447$ ) en zone indemne de FCM (continent français)**



### 1.1.3. Receiver-operating characteristic (ROC) analyse

#### 1.2.2.1. Généralités

L'analyse ROC, depuis son introduction à la fin des années 60 dans le domaine des sciences médicales, fait partie des outils standards permettant d'évaluer un test de diagnostic clinique. Cette technique fait partie de celles recommandées par le manuel des normes pour les tests de diagnostic et les vaccins de l'OIE [OIE, 2002] pour la détermination de la valeur-seuil d'un test de diagnostic.

Les conditions suivantes sont indispensables pour pouvoir conduire ce type d'analyse :

- Le statut de l'animal vis-à-vis de la maladie ou de l'infection doit être connu. Un test de référence avec des qualités intrinsèques proches du « gold standard » est nécessaire.
- Le test doit rendre un résultat de type quantitatif et non binaire.

Pour un test quantitatif, l'estimation des paramètres de sensibilité et de spécificité dépend de la valeur-seuil (ou cut-off) retenue. En effet, pour chaque cut-off possible du test correspond un tableau de contingence (tableau I) classant les animaux en 4 catégories : vrais-positifs (VP), faux-positifs (FP), vrais-négatifs (VN) et faux-négatifs (FN). L'utilisation de l'analyse ROC permet de représenter graphiquement, sur une unité de surface, les différents couples de sensibilité et de spécificité calculées pour chaque valeur-seuil possible, c'est-à-dire la performance intrinsèque du test associée à chaque cut-off. Le graphique (figure 16) a pour abscisse (1-spécificité) c'est-à-dire la fraction de faux-positifs et en ordonnée la sensibilité (se) c'est-à-dire la fraction de vrais-positifs.

Toutes les combinaisons possibles de se/sp peuvent alors se résumer grâce à un seul paramètre : l'aire sous la courbe (AUC). Celle-ci varie de 0 à 1. La courbe obtenue avec un test de diagnostic permettant une discrimination parfaite entre les échantillons négatifs et positifs (absence de superposition entre les deux groupes) passe par les coordonnées {0 ;1}. On est alors dans le cas du gold standard où la sensibilité et la spécificité sont de 100 %. L'aire sous la courbe permet ainsi d'évaluer le **pouvoir discriminant du test**. On peut alors classer les tests en fonction de leur niveau de performance : les tests non informatifs ( $AUC = 0,5$ ), les tests peu performants ( $0,5 < AUC \leq 0,7$ ), les tests modérément performants ( $0,7 < AUC \leq 0,9$ ) et les tests fortement performants ( $AUC > 0,9$ ) [Greiner *et al.*, 2000].

Cette donnée peut également être utilisée pour comparer des tests.

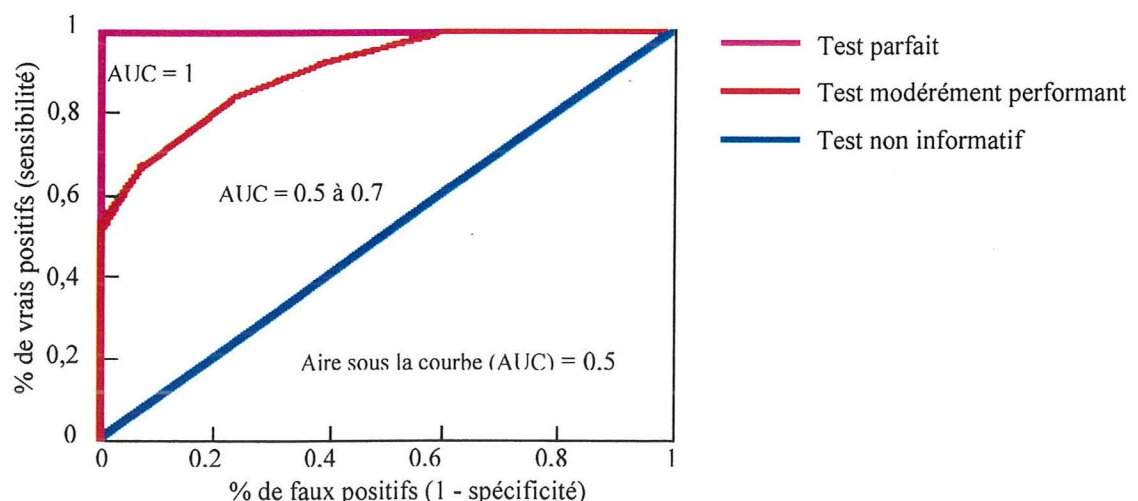
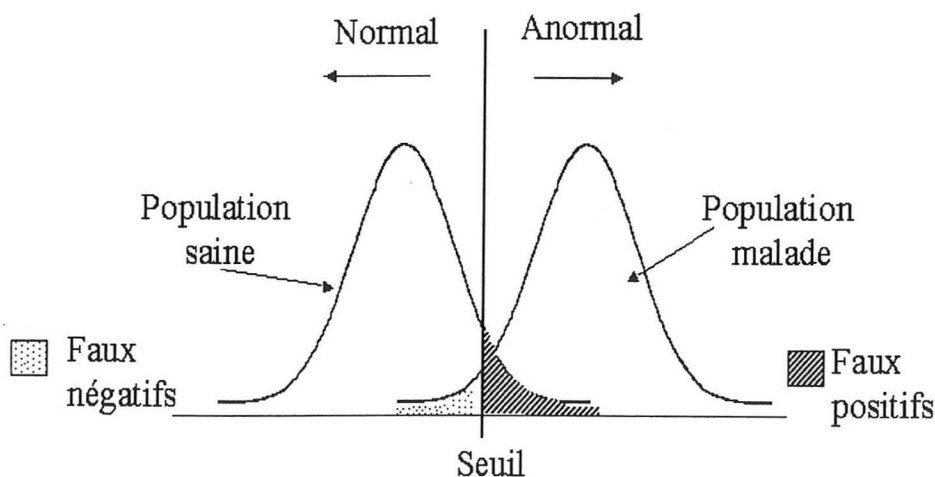


Figure 16 : Différentes allures de courbe ROC en fonction du pouvoir discriminant du test

Un autre atout de l'analyse ROC est de permettre la **sélection de la valeur-seuil** la plus pertinente en fonction des objectifs du test et de la situation épidémiologique de la maladie dans la population cible. La méthode communément utilisée pour choisir la valeur déterminant le classement des animaux en positifs ou négatifs est la méthode de la distribution gaussienne (normale). Le cut-off est alors défini comme la moyenne plus deux écarts-types de l'échantillon négatif de référence. 97,5 % de la population négative se trouvent alors en deçà de la valeur-seuil. On atteint alors une spécificité de 97,5 %. Cette méthode trouve ses limites dans le fait que la distribution des valeurs du test au sein d'une population n'est pas toujours normale. De plus et surtout, on ne considère à aucun moment la sensibilité du test.

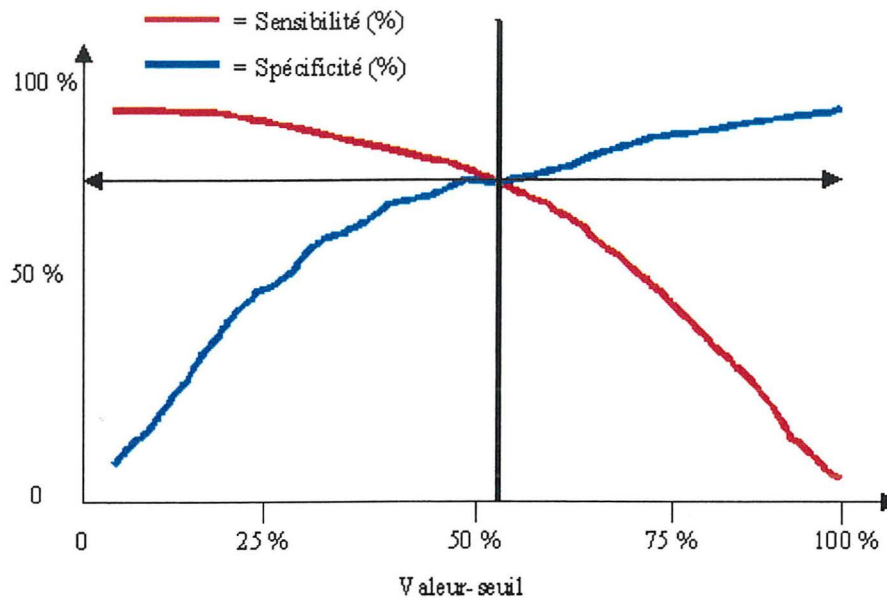
La connaissance du couple sensibilité / spécificité pour chaque valeur-seuil possible du test permet de choisir cette dernière en fonction des valeurs de sensibilité et de spécificité recherchées. On peut ainsi introduire la notion de coûts des faux-positifs et/ou des faux-négatifs. En fonction de la maladie étudiée, de sa prévalence, des conséquences économiques et sanitaires possibles, le poids lié aux faux-positifs et aux faux-négatifs sera différent. Si la zone surveillée est indemne de la maladie étudiée, l'obtention de faux-positifs pourra avoir des conséquences importantes, notamment s'il s'agit d'une maladie réputée contagieuse pouvant entraîner des décisions préfectorales, nationales voire européennes conduisant notamment à l'arrêt des échanges commerciaux des espèces sensibles (ruminants). Dans ce cas de figure, on recherchera un test très spécifique. On retiendra alors un seuil plus élevé (figure 17). A l'inverse, si la prévalence de la maladie est forte, on s'orientera vers une très bonne sensibilité du test (diminution du seuil). On a ainsi la possibilité de moduler les performances du test suivant le contexte épidémiologique et la population cible.

Toutefois, afin de ne pas choisir arbitrairement la valeur-seuil optimale dans un contexte donné, il existe un panel d'approches permettant la sélection du cut-off en fonction des objectifs fixés [Greiner *et al.*, 2000]. L'approche la plus simple consiste à postuler que le coût des faux-positifs et des faux-négatifs sont identiques et que la prévalence dans la population cible est de 50 %. Le point de la courbe le plus en haut à gauche (figure 16) optimise alors la sensibilité et la spécificité du test ( $se = sp$ ). L'approche à l'aide de deux courbes ROC, TG-ROC (Two Graph – ROC), permet de visualiser plus rapidement le cut-off à sélectionner [Greiner, 1995]. Elle constitue à représenter sur un même graphique (figure 18) à la fois la courbe de sensibilité et de spécificité du test en fonction de la valeur-seuil. Le cut-off pour lequel la sensibilité et la spécificité sont identiques est alors obtenu au point d'intersection des deux courbes.



**Figure 17 : Représentation graphique de l'effet de la variation du seuil sur la proportion des faux-positifs et des faux-négatifs**





**Figure 18 : Exemple de courbe TG-ROC [Greiner, 1995 & 1996]**

La connaissance de la qualité intrinsèque du test permet également une meilleure **interprétation des résultats individuels**. En effet, l'interprétation d'un résultat individuel dépend de la valeur prédictive du test. Or, cette dernière est étroitement liée à la sensibilité, la spécificité et la prévalence de la maladie (figure 20). Si la prévalence de la maladie dans la population cible est proche de zéro, la valeur prédictive positive (VPP) du test sera faible. Le choix d'un test très spécifique (sp proche de 1) est alors nécessaire. La connaissance de la sensibilité et de la spécificité permet ainsi d'estimer ces valeurs prédictives en fonction de la prévalence de la maladie.

La sensibilité et la spécificité interviennent également pour le calcul de la prévalence réelle (Pr) de la maladie qui peut être différente de la prévalence observée c'est-à-dire de la prévalence apparente (Pa). Elles diffèrent en effet par les erreurs liés aux faux- positifs et faux-négatifs (figure 19).

$$Pr = \frac{Pa + Sp - 1}{Se + Sp - 1}$$

**Figure 19 : Formule de la prévalence réelle d'une maladie, [ Rogan & Gladen, 1978]**

- valeur prédictive positive (VPP) = probabilité que l'animal soit malade alors que le test est positif
- valeur prédictive négative (VPN) = probabilité que l'animal ne soit pas malade alors que le test est négatif

$$VPP = \frac{Se * P}{Se * P + (1 - Sp) * (1 - Pr)}$$

$$VPN = \frac{Sp * (1 - P)}{Sp * (1 - P) + (1 - Se) * P}$$

**Figure 20 : Définitions des valeurs prédictives positives (VPP) et négatives (VPN) (d'après le théorème de Bayes)**

#### 1.2.2.2. Démarche suivie

Le statut des animaux vis-à-vis de l'infection est défini par rapport au résultat du test de référence choisi, la RT-PCR.

L'analyse est développée successivement pour l'échantillon de 64 sérums ( $n_1 = 64$ ) de l'année 2000 (test cELISA I) et pour l'échantillon  $n_2$  ( $n_2 = 338$ ) de l'année 2001 (test cELISA II).

Pour l'échantillon  $n_2$ , plusieurs courbes ROC ont été dressées.

- Dans un premier temps, seuls les animaux non vaccinés ( $n_2' = 104$ ) ont été conservés (retrait des animaux vaccinés ou au statut vaccinal inconnu).
- Dans un second temps, afin de définir clairement les conditions dans lesquelles les animaux ont été prélevés, nous nous sommes restreints aux animaux non vaccinés dont le motif de prélèvement était la suspicion clinique ( $n_2'' = 86$ ). Nous sommes ainsi dans les conditions d'un sérodiagnostic.
- Afin de restreindre la possibilité que des animaux aient été en contact avec le virus en 2000 (infection inapparente), seuls les animaux non vaccinés de l'espèce ovine ont été maintenus dans l'échantillon ( $n_2''' = 88$ ). Il s'agit essentiellement d'agnelles de l'année. Le motif de prélèvement a également été pris en compte ( $n_2'''' = 77$ ).
- Une dernière approche a été d'ajouter à l'échantillon d'animaux corses des animaux contrôlés sur le continent. Leur statut vis-à-vis de l'infection est connu puisque la zone est indemne de maladie (animaux non malades). Ainsi, les résultats des bovins du continent analysés en 2002 ont complété l'échantillon soit 447 résultats.

#### 1.2.2.3. Logiciels utilisés

Les données ont été traitées statistiquement à l'aide du logiciel STATA 7.0 <sup>ND</sup> (<http://www.stata.com/>).

Les graphiques représentant la distribution des valeurs de pourcentage d'inhibition ont été édités grâce à la commande « graph ». Les commandes « roctab » et « rocfits » ont été utilisées pour la construction des courbes ROC. Les statistiques de dispersion ont été obtenues en utilisant la commande « tabstat ».

Les intervalles de confiance des valeurs de sensibilité et de spécificité ont été calculés grâce au logiciel WinEpiscop 2.0 [Thrusfield *et al*, 2001] ([http://epiweb.massey.ac.nz/Epidemiological\\_analysis\\_software.htm](http://epiweb.massey.ac.nz/Epidemiological_analysis_software.htm)).

La courbe TG-ROC a été construite à partir du logiciel TG-ROC (<http://city.vetmed.fu-berlin.de/~mgreiner/TG-ROC/zipENGLISH97/>). Elle permet la sélection du cut-off optimisant les valeurs de sensibilité et de spécificité du test (cELISA I et cELISA II).



## 1.2. Comparaison des deux tests cELISA

### 1.2.1. Principe général de l'étude

En 2000, un premier test ELISA de compétition (cELISA I) a été retenu par le laboratoire national de référence sur la fièvre catarrhale du mouton (CIRAD-emvt, Montpellier). Compte tenu des problèmes de répétabilité et de paramétrage du kit, un autre test ELISA de compétition (cELISA II) lui a été préféré à partir de 2001. Il avait notamment été relevé lors du début de la campagne de surveillance sur le continent en 2001 qu'un certain nombre de bovins du continent (zone indemne) se révélait positif.

La démarche suivie a été de reconstruire avec le test cELISA II plusieurs sérums provenant d'ovins prélevés en 2000 et conservés en micro-plaques à  $-18^{\circ}\text{C}$ . L'ensemble des résultats ont ensuite été analysés statistiquement afin d'évaluer la concordance des résultats obtenus avec les deux tests.

### 1.2.2. Données

#### 1.2.2.1. Généralités

La totalité des sérums prélevés en 2000 ont été conservés en micro-plaques soit 8825 sérums. Ils se répartissent en trois catégories de résultats : 2 442 positifs, 1 846 douteux et 4 537 négatifs (tableau III), provenant de 206 élevages.

**Tableau III : Distribution des sérums prélevés en 2000 selon le résultat sérologique (cELISA I)**

		résultat sérologique (cELISA I)			
		négatif	douteux	positif	Total
espèce	BOVINS	2 095 (45,2 %)	1 198 (25,8 %)	1 344 (29,0 %)	4 637
	CAPRINS	19	119	11	149
	OVINS	4 537 (51,4 %)	1 846 (20,9 %)	2 442 (27,7 %)	8 825
Total		6 651	3 163	3 797	13 611

#### 1.2.2.2. Sélection des sérums

Dans un premier temps, deux plaques ont été sélectionnées au hasard parmi les élevages déclarés infectés en 2000 (éleveurs 1 et 9) et une plaque parmi les élevages non officiellement déclarés infectés (élevage 2), soit au total 241 sérums.

Au vu des premiers résultats, l'échantillon a été élargi avec 165 sérums supplémentaires (choix aléatoire de 3 plaques) provenant de 6 autres élevages dont 4 ont été déclarés « foyer de fièvre catarrhale du mouton » en 2000 (tableau IV).

**Tableau IV :Détail concernant les sérums de 2000 recontrôlés par le test cELISA II**

		résultat sérologique (cELISA I)			
		négatif	douteux	positif	Total
Élevages	Élevage 1	12	35	31	78
	Élevage 2	24	46	7	77
	Élevage 3	1	6	3	10
	Élevage 4	2	19	16	37
	Élevage 5	9	27	42	78
	Élevage 6	0	13	25	38
	Élevage 7	0	0	1	1
	Élevage 8	0	0	1	1
	Élevage 9	3	22	61	86
Total		51 (12,6 %)	168 (41,4 %)	187 (46,0 %)	<b>406</b>

Un total de 406 sérums a ainsi fait l'objet d'un test de concordance en utilisant le test kappa.

### 1.2.3. Coefficient kappa

Ce coefficient est communément utilisé pour évaluer la variabilité des résultats entre deux observateurs (ou bien celle d'un seul observateur à deux moments différents) pour un diagnostic ou une observation. La seule mesure du pourcentage des observations qui concordent ne donne en effet pas une idée juste du degré d'accord entre les observateurs puisqu'il persiste une part de hasard. Le coefficient kappa permet d'évaluer le degré de concordance entre les observateurs en tenant compte du hasard. « La concordance est la conformité, la similitude de deux ou plusieurs jugements ou informations de même nature qui se rapportent au même objet » (d'après Grenier, Masson, 1996). La concordance se distingue donc de la corrélation dans le sens où elle implique appariement et jugements de même nature.

Le coefficient kappa s'applique plus généralement à des variables discontinues. Ses valeurs sont comprises entre -1 et 1. Un kappa proche de -1 signifie une discordance complète. Dans le cas d'un jugement qualitatif, une échelle d'évaluation de la qualité de la concordance a été proposé par Brennan & Hays (1992) : entre 0 et 0,20, la concordance est mauvaise, entre 0,21 et 0,40, la concordance est faible, entre 0,61 et 0,80, la concordance est bonne et entre 0,81 et 1, la concordance est excellente.

Ce coefficient présente certaines limites [Briand]. Il est notamment sensible au déséquilibre de fréquence. L'estimation du coefficient kappa est par exemple plus élevée lorsque les effectifs marginaux sont répartis de façon équilibrée sur les différentes modalités de réponse que lorsqu'ils présentent un fort déséquilibre en faveur d'une même catégorie de résultat. De plus, il convient de comparer des tests ayant les mêmes classes de résultats.

Le coefficient kappa a été calculé avant et après réévaluation du seuil du test cELISA I par l'analyse ROC. La zone douteuse n'a pas été maintenue afin de se conformer aux recommandations du fabricant et de ne conserver que deux catégories d'animaux pour les deux tests : animaux malades et non malades. Cette simplification permet de se ramener à comparer des tests ayant un même nombre de classes de résultats de même amplitude (0 à 49,9 % et 50 % à 100 %).



## 2. Résultats

### 2.1. Évaluation du test de diagnostic cELISA I

En prenant la RT-PCR comme test de référence, l'échantillon de 64 sérums se compose de 31 animaux non malades et de 33 animaux malades.

La courbe obtenue avec l'analyse ROC est représentée par la figure 21. La valeur-seuil correspondant aux meilleures valeurs de sensibilité et de spécificité est à 70 % d'inhibition. Les sensibilité et spécificité associées sont alors respectivement de 84,9 % (IC (95 %) : [72,6 - 97,1]) et de 83,9 % (IC (95 %) : [70,9 - 96,8]).

A la valeur-seuil préconisée par le fabricant, c'est-à-dire 50 % d'inhibition, ces valeurs sont respectivement de 93,9 % (IC (95 %) : [85,8 % - 100 %]) et de 51,6 % (IC (95 %) : [34,0 - 69,2]).

L'aire sous la courbe est de 0,84 avec un intervalle de confiance de [0,73 - 0,95].

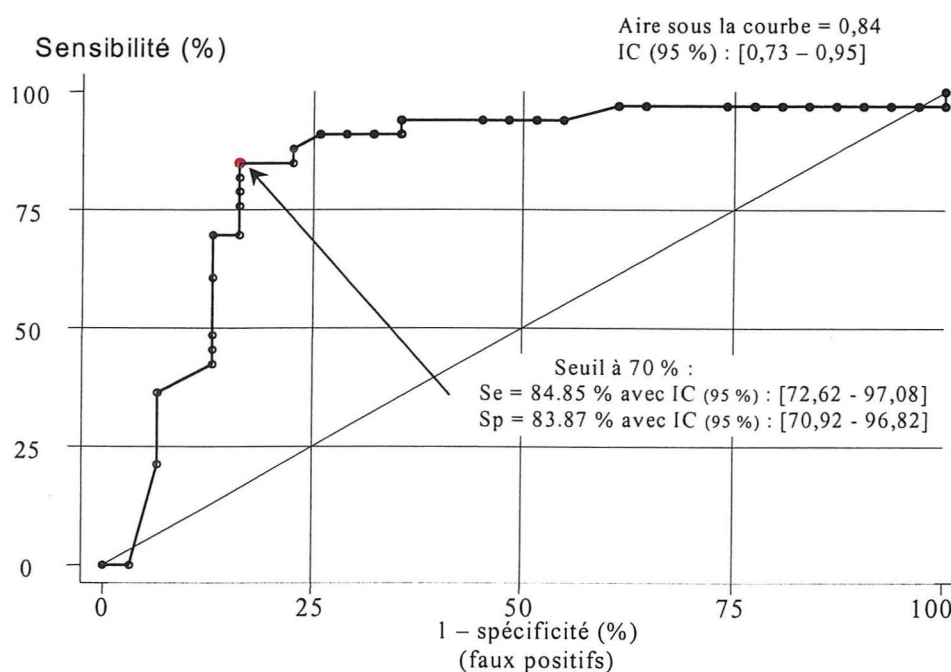


Figure 21 : Courbe ROC obtenue avec le test cELISA I ( $n_1 = 64$ )

## 2.2. Évaluation du test de diagnostic cELISA II

Afin de déterminer la capacité du test cELISA II à discriminer les animaux infectés naturellement (virus sauvage) des animaux non infectés, seuls les animaux non vaccinés de l'échantillon  $n_2$  ont été retenus soit 103 animaux. Parmi ces sérums, 41 se sont révélés négatifs en RT-PCR soit 41 sérums sont considérés provenir d'animaux non malades. A l'inverse, 62 ont un résultat positif en RT-PCR et constituent notre pool d'animaux malades.

La valeur globale du test peut être exprimée par l'aire sous la courbe issue de l'analyse ROC. Ainsi, en considérant l'ensemble des espèces pour lesquelles on dispose des résultats sérologique et virologique, l'aire sous la courbe est de 0,70 avec un intervalle de confiance ( $\alpha=0,05$ ) compris entre 0,58 et 0,81.

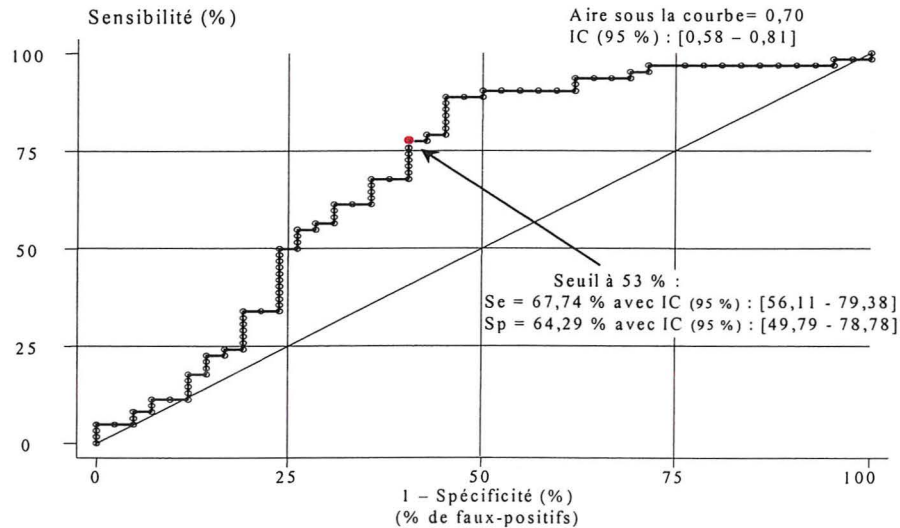


Figure 22 : Courbe ROC obtenue en retenant seulement les sérums provenant d'animaux non vaccinés ( $n_2' = 103$ , cELISA II)

L'information sur le motif de prélèvement a permis de mieux définir les caractéristiques de l'échantillon. En conservant uniquement les animaux ayant fait l'objet d'une suspicion clinique de FCM, on obtient la courbe ROC de la figure 19. L'aire sous la courbe est alors de 0,78 (avec IC 95 % = [0,68 - 0,89]).

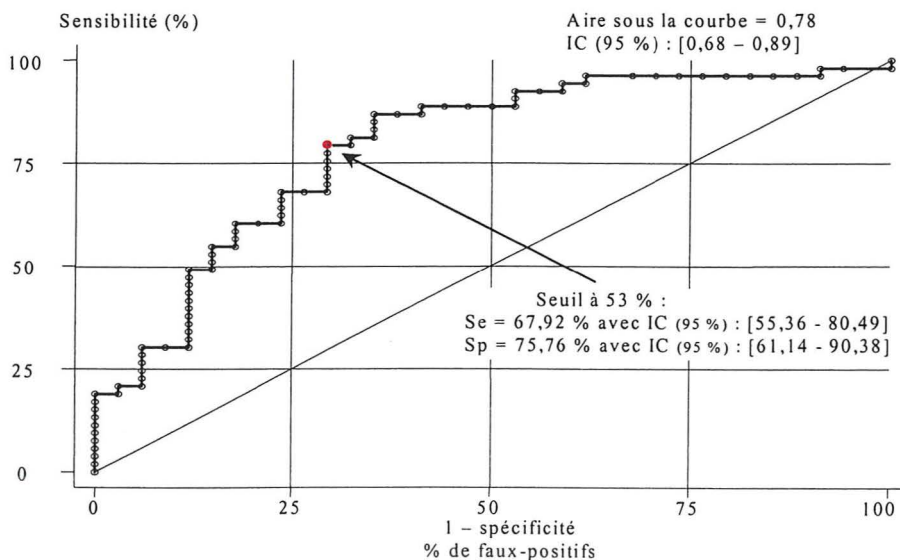
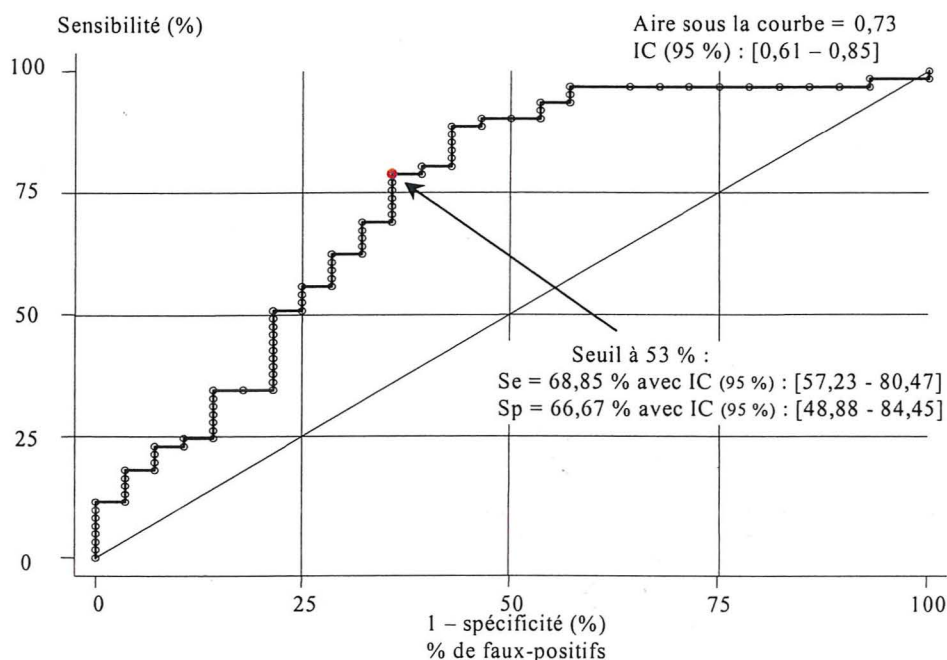


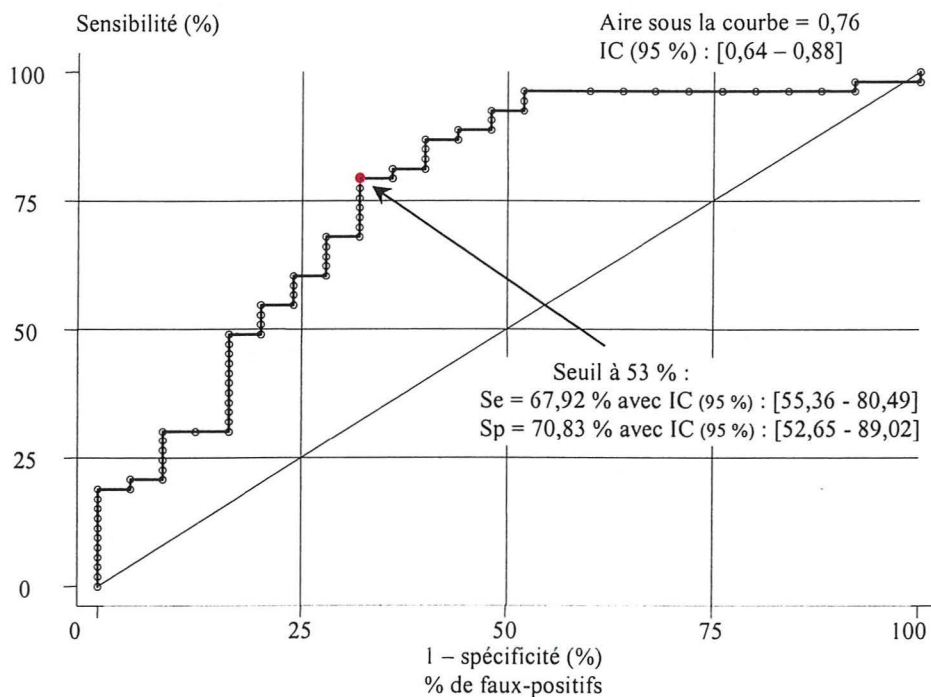
Figure 23 : Courbe ROC obtenue en conservant dans l'échantillon que les animaux non vaccinés dont le motif de prélèvement était la suspicion clinique ( $n_2'' = 85$ , cELISA II)



L'analyse a été poursuivie en se restreignant à la seule espèce ovine. On souhaite ainsi limiter le risque d'inclure dans l'échantillon un animal qui a pu rencontrer le virus en 2000 mais ne pas exprimer de signes cliniques (cas des caprins, bovins). Les résultats sont présentés dans les figures 24 & 25. Dans le premier cas, le facteur « motif de prélèvement » n'est pas retenu ( $n = 88$ ). L'aire sous la courbe (AUC) est ainsi de 0,73 (IC 95 % = [0,61 – 0,85]). Dans le second cas, seuls les animaux cliniquement suspects ont fait l'objet de l'analyse. L'AUC correspondante est de 0,76 (IC 95 % = [0,64 – 0,88]).

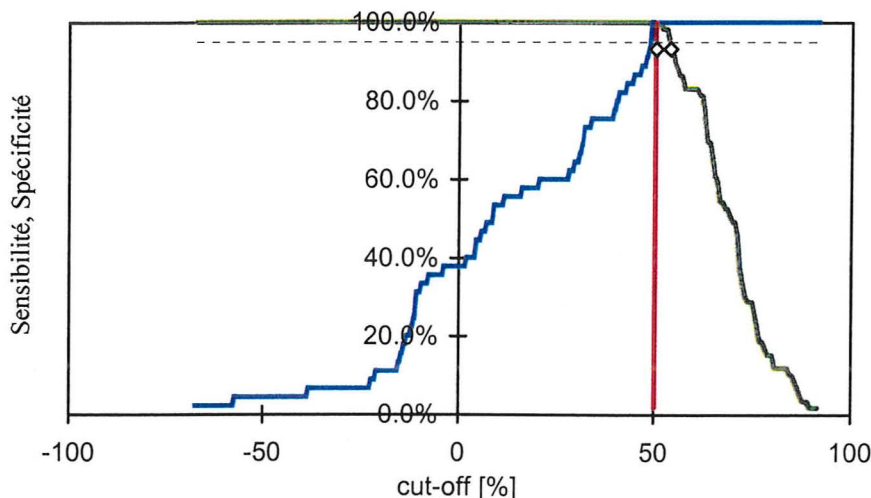


**Figure 24 :** Courbe ROC obtenue en retenant que les sérums provenant d'ovins non vaccinés ( $n_2''=87$ , cELISA II)



**Figure 25 :** Courbe ROC obtenue en retenant que les sérums provenant d'ovins non vaccinés dont le motif de prélèvement était une suspicion clinique de fièvre catarrhale du mouton ( $n_2''' = 76$ , test cELISA II)

Pour chacune de ces analyses, le détail des résultats (Cf. annexe V) conduit à définir la valeur de 53 % comme le cut-off pour lequel la sensibilité et la spécificité du test pour cette population et dans ces conditions particulières d'utilisation sont les meilleures (figure 26). La valeur de sensibilité varie entre 68 et 70 % avec des intervalles de confiance pour  $\alpha = 0,05$  compris entre 55 et 80 %. Suivant les scénarios, la spécificité varie de 64 % (sérums tout venant) à 76 % (sérums avec une suspicion clinique) avec des intervalles de confiance pour  $\alpha = 0,05$  ayant un minimum à 49 % et un maximum à 90 %. En cas de suspicion clinique, une spécificité de 100 % est atteinte pour une valeur-seuil de 72 %. La sensibilité associée est alors de 19 % (IC 95 % = [8 % – 29 %]).

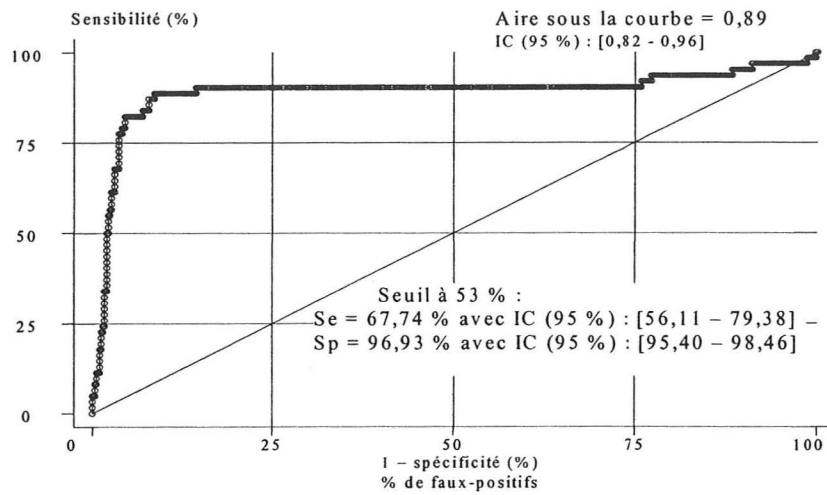


**Figure 26 : Courbe TG-ROC permettant de mettre en évidence le cut-off correspondant aux meilleures valeurs de sensibilité et de spécificité, le poids de ces deux paramètres étant considéré comme équivalent.**

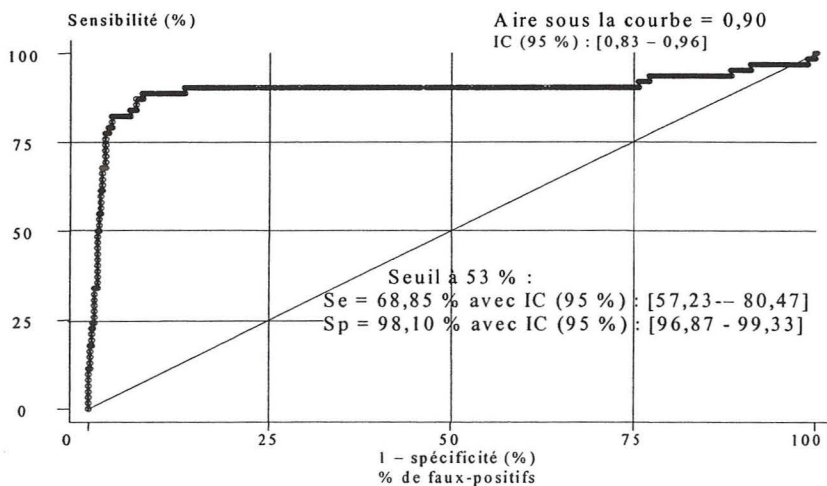
En incorporant dans l'échantillon les résultats disponibles dans une zone indemne (surveillance sérologique par sondage sur le continent français), on peut améliorer la précision et la valeur de la spécificité. Dans un premier temps, seuls les résultats bovins de l'année 2002 ont été ajoutés soit 447 données. On obtient alors les courbes ROC des figures 26 & 27. Au seuil de 53 %, la sensibilité ne varie pas. Par contre, la spécificité associée atteint 96,9 % (IC à 95 % = [95,4 – 98,5]). Dans ce cas, le motif de prélèvement n'est pas pris en compte (figure 26). Elle est de 98,10 % (IC à 95 % = [96,9 – 99,3]) si seuls les animaux corses cliniquement suspects sont intégrés (figure 27). Les aires sous la courbe sont respectivement de 0,89 (IC à 95 % = [0,82 – 0,96]) et de 0,90 (IC à 95 % = [0,87 – 0,98]).

Puis, tous les résultats bovins de zone indemne de fièvre catarrhale du mouton ont été intégrés soit un total de 3755. La spécificité calculée au seuil de 53 % est de 99,6 % (IC à 95 % = [99,4 – 99,8]). L'aire sous la courbe atteint 0,92 (IC à 95 % = [0,87 – 0,98]).

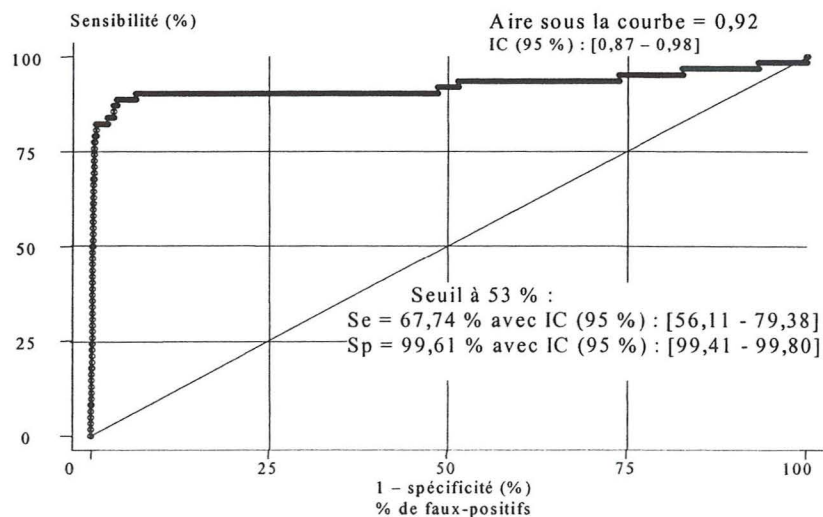




**Figure 27 :** Courbe ROC obtenue en ajoutant les 447 résultats de bovins provenant d'une zone indemne de FCM (continent français, campagne 2002) au 103 sérums corses ( $n_2'''=550$ , cELISA II)



**Figure 28 :** Courbe ROC obtenue en ajoutant les 447 résultats de bovins provenant d'une zone indemne de FCM (continent français, campagne 2002) aux 87 sérums ovins corses ( $n=534$ , cELISA II)



**Figure 29 :** Courbe ROC obtenue en ajoutant les 3 755 résultats de bovins provenant d'une zone indemne de FCM (continent français, campagne 2001 et 2002) aux 103 sérums corses ( $n = 3858$ , cELISA II)

	« Gold standard »	Nb malades	Nb non malades	Aire sous la courbe (AUC)	Intervalle de Confiance à 95 %	Sensibilité au seuil de 53 %	Spécificité au seuil de 53 %	Commentaires
1	RT-PCR	62	41	0,70	0,73 – 0,95	67,7 %	64,3 %	Animaux non vaccinés
2	RT-PCR	53	32	0,78	0,68 – 0,89	67,9 %	<b>75,8 %</b>	Non vaccinés + suspicion clinique
3	RT-PCR	61	26	0,73	0,61 – 0,85	<b>68,9 %</b>	66,7 %	Ovins non vaccinés
4	RT-PCR	53	23	0,76	0,64 – 0,88	67,9 %	70,8 %	Ovins non vaccinés + suspicion clinique
5	RT-PCR	62	488	0,89	0,82 – 0,96	67,7 %	<b>96,9 %</b>	n°1 + bovins du continent (2002)
6	RT-PCR	61	473	0,90	0,83 – 0,96	68,9 %	<b>98,1 %</b>	n°2 + bovins du continent (2002)
7	RT-PCR	62	3796	0,92	0,87 – 0,98	67,7 %	<b>99,6 %</b>	n°1 + bovins du continent (2001 & 2002)

**Tableau V : Synthèse de l'ensemble des résultats de l'analyse ROC en fonction des critères de sélection des sérums**

Les approches les plus pertinentes au vu de l'échantillon de 2001 disponibles, soit 338 sérums, correspondent aux analyses travaillant sur les sérums de l'échantillon provenant d'animaux non vaccinés dont le motif de prélèvement était la suspicion clinique avec introduction ou non du facteur « espèce » (scénarios 2 & 4 ; tableau V). Lorsque toutes les espèces sont prises indifféremment, les valeurs de sensibilité et de spécificité obtenues pour le test cELISA II sont respectivement de 67,9 % (IC à 95 % = [55,3 % - 80,5 %]) et de 75,8 % (IC à 95 % : [61,1 % - 90,4 %]) pour un total de 85 sérums. L'aire sous la courbe est de 0,78 (IC à 95 % = [0,68 - 0,89]). Si seuls les ovins sont conservés, la spécificité baisse légèrement en valeur absolue mais cette différence n'est pas significative (76 sérums). La spécificité est alors de 70,8 % (IC à 95 % : [52,7 – 89,0 %]). L'aire sous la courbe n'est pas non plus significativement différente à 0,76 (IC à 95 % : [0,64- 0,88]).

L'ajout des résultats bovins provenant d'une zone indemne du virus permet d'augmenter de manière importante le nombre d'animaux non malades et non infectés et ainsi observer le comportement de la spécificité au seuil de 53 %. En admettant comme « malades » les animaux virologiquement positifs suite à une suspicion clinique, la spécificité atteint 98,1 % (scénario 6).



## 2.3. Comparaison des deux tests

### 2.3.1. Sans réajustement du seuil

Les valeurs de pourcentage d'inhibition obtenues avec le test cELISA I et cELISA II se distribuent respectivement autour la moyenne de 54,4 % ( $\sigma^2 = 10,0$ ) et de 20,6 ( $\sigma^2 = 12,5$ ). Ces moyennes sont significativement différentes ( $t = 51,21$  et  $p < 0,001$ ).

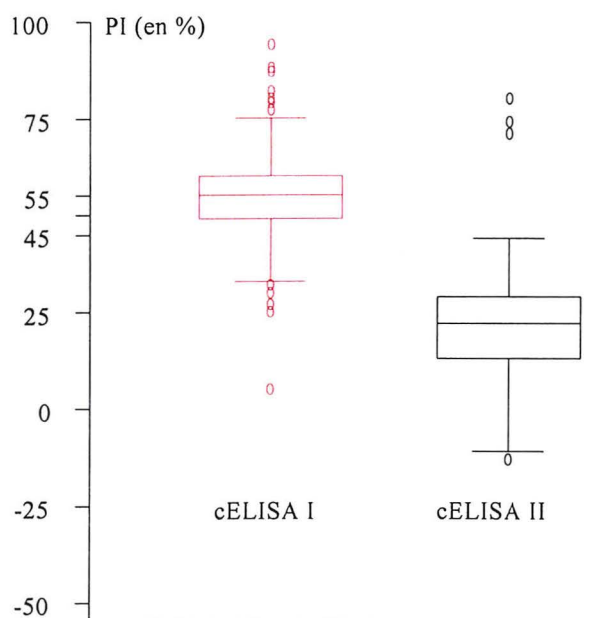


Figure 30 : Boîte à moustaches représentant la dispersion des pourcentages d'inhibition (PI) , n=406, en fonction du test utilisé (cELISA I vs cELISA II)

Tableau VI : Résultats obtenus avec les sérums prélevés en 2000 (n=406) selon le test cELISA utilisé, les deux tests ayant une valeur-seuil fixée à 50 %.

		cELISA I seuil à 50 %		
		négatif	positif	Total
cELISA II	négatif	114	289	403
	positif	0	3	3
Total		114	292	406

Le coefficient kappa calculé est de 0,006 ( $\sigma^2 = 0,005$ ) lorsque la valeur-seuil des deux tests est maintenue à 50 %. 289 résultats sont discordants soit 71,2 % (Cf. tableau V).

### 2.3.2. Avec réajustement du seuil

Le seuil du test cELISA I a été redéfini au vu des résultats de l'analyse ROC (Cf. 2.1.) et des valeurs de pourcentage d'inhibition obtenues pour les contrôles négatifs lors des analyses. Ils se révélaient en effet globalement supérieurs à la limite de 25 % fixé comme critère de validité par le fabricant. Le seuil du test cELISA est ainsi maintenant fixé à 70 %. Le cut-off du test cELISA II est maintenu à 50 % au vu des résultats de l'analyse ROC (Cf. 2.2.)

Le nouveau coefficient kappa calculé est de 0,190 ( $\sigma^2 = 0,035$ ). Le nombre de résultats discordants est alors de 16 sur 406 soit 3,9 % (Cf. tableau VI).

**Tableau VII : Résultats obtenus avec les sérums prélevés en 2000 (n = 406) selon le test cELISA utilisé, le test cELISA I ayant une valeur-seuil fixée à 70 %.**

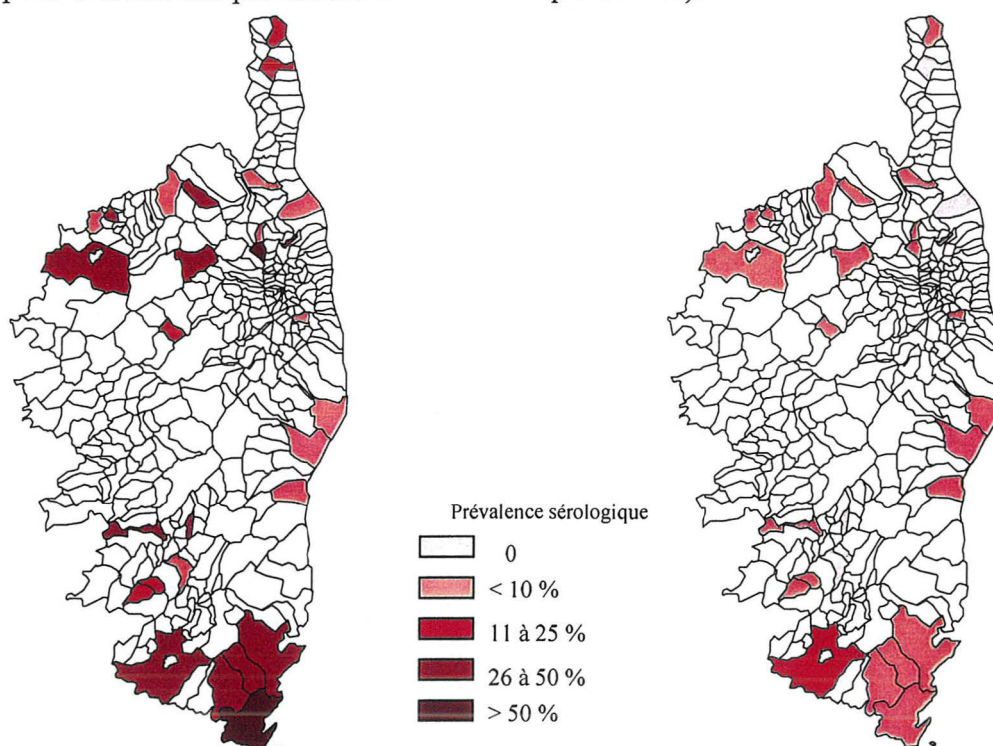
		cELISA I seuil à 70 %		Total
cELISA II		négatif	positif	
	négatif	388	15	403
	positif	1	2	3
Total		114	292	406

#### 2.4. Exemples d'applications : calcul de la prévalence apparente chez les ovins en 2000

En reprenant toutes les données sérologiques de l'année 2000 de l'espèce ovine, on peut dresser une carte représentant les prévalences sérologiques apparente en fonction de la situation géographique (communes corses).

La figure 27 fait apparaître la prévalence apparente lorsque le test cELISA à une valeur-seuil fixée à 50 %. La figure 28 correspond à la prévalence apparente lorsque le test cELISA à une valeur-seuil fixée à 70 %.

Cette correction fondée sur les résultats de l'analyse ROC conduit à revoir à la baisse la prévalence apparente de la fièvre catarrhale du mouton chez les ovins en 2000. Dans les communes pour lesquelles on dispose de données sérologiques suffisantes pour calculer une prévalence, celle-ci atteint dans la majorité des communes un pourcentage inférieur à 10. La commune de Sartène fait exception avec une prévalence observée de 21 % (+/- 8 %). Quatre communes (Borgho, Luri, Moca croce, Zevaco) voient leur prévalence observée ramenée à zéro (prévalence établie à partir de l'effectif total ou d'un échantillon supérieur à 149 c'est-à-dire suffisant pour détecter une prévalence de 2 % au risque de 5 %).



**Figure 31 : Prévalence apparente chez les ovins en 2000 avec un seuil défini à 50 % pour le test cELISA I**

**Figure 32 : Prévalence apparente chez les ovins en 2000 avec un seuil défini à 70 % pour le test cELISA I**



### 3. Discussion

Le test cELISA est évalué dans des conditions particulières qu'il convient de rappeler. Les données à la base de cette étude ont été collectées en période d'épizootie de fièvre catarrhale du mouton en Corse. Le test sérologique était utilisé à des fins de diagnostic rapide. L'objectif, pendant cette période, était de détecter tous les foyers de bluetongue chez les ovins. Il s'agit en effet d'une maladie réputée contagieuse pour cette espèce. A ce titre, elle fait l'objet d'une réglementation propre qui implique la mise en place de mesures de police sanitaire dans le cas d'une confirmation de la circulation du virus. Les prises de sang ont de ce fait été réalisées dès les premiers signes cliniques évocateurs de bluetongue. Cet élément sera d'autant plus important à prendre en compte dans la discussion des résultats que le test de référence choisi est un test virologique qui par nature détecte le virus.

Le premier point de discussion porte ainsi sur le choix du test de référence.

La RT-PCR est réputée pour ses très bonnes valeurs de sensibilité et de spécificité ce qui, de ce point de vue, en fait un très bon test de référence et permet de le proposer comme « gold standard ». Le principal problème des techniques PCR réside en général dans les risques de contamination de laboratoire qui peuvent conduire à des résultats faussement positifs. C'est pourquoi les laboratoires réalisant ce type d'analyses doivent respecter certaines règles afin de minimiser au maximum ce risque. Ce test n'est ainsi pas un test parfait. Ces valeurs de sensibilité et de spécificité ne sont pas également à 100 % même si elles s'en rapprochent lorsqu'il est utilisé dans de très bonnes conditions. Cependant, comme pour beaucoup de tests, ces valeurs ne sont pas bien documentées.

La principale limite du choix de ce test provient du fait que les deux tests ne détectent pas la même entité biologique. La RT-PCR met en évidence l'éventuelle infection par l'orbivirus en amplifiant un brin d'ARN spécifique du virus alors que l'ELISA de compétition révèle l'éventuelle présence d'anticorps. Or, la phase de virémie étant aussi antérieure à l'apparition des anticorps, il existe ainsi une période pendant laquelle un animal déclaré infecté au vu du résultat virologique pourra obtenir un résultat sérologique négatif (figure 3). Ce résultat en apparence faussement négatif n'est cependant pas attribuable à un défaut de sensibilité du test. Dans notre étude, ce cas de figure est d'autant plus probable que la plupart des prélèvements de sang proviennent d'animaux en début de phase clinique (suspicion clinique).

Lorsque le motif de prélèvement n'est pas pris en compte dans l'analyse ROC, l'échantillon contient des animaux prélevés dans le cadre d'un suivi d'élevage reconnu infecté c'est-à-dire des animaux *a priori* sans signes cliniques de la maladie. Les animaux survivants sont ainsi potentiellement des animaux qui, lors du passage du virus, se sont naturellement immunisés. Dans ce schéma, la RT-PCR peut révéler un animal négatif (animal prélevés après la phase de virémie) alors que le résultat ELISA peut attester de la présence d'anticorps spécifiques de la fièvre catarrhale du mouton. La démarche conclura alors à un résultat faussement positif sans que la spécificité du test n'en soit la cause. C'est pourquoi, dans notre étude, ces cas potentiels ont été éliminés en conservant dans l'analyse que les animaux dont le motif de prélèvement était la suspicion clinique.



En zone infectée, sur une population de ruminants sensibles à la fièvre catarrhale du mouton, lorsque le test cELISA est utilisé à des fins de diagnostic précoce de l'infection, le test II obtient ses meilleures valeurs de sensibilité et de spécificité au seuil de 53 %. Ce seuil reste identique quelque soit le scénario retenu. Pour ce kit, le fabricant a validé la valeur de 50 % d'inhibition pour séparer la population de malades des non malades. Le laboratoire national de référence français en matière de FCM a de plus délimité une zone douteuse comprise entre 45 et 55 %. Le résultat obtenu valide ainsi la valeur-seuil utilisée. *A posteriori*, les données provenant de la population indemne du continent français confirment également la pertinence de cette valeur-seuil. En effet, la moyenne du pourcentage d'inhibition obtenue sur cette population indemne est de 18 % avec un écart type de 12 %. En fixant le seuil à la valeur correspondant à la moyenne plus trois écarts types et compte tenu du fait que la distribution suit une loi normale, on atteint une spécificité de 99, 5 %. Ainsi, avec cette méthode, à la valeur seuil de 54 %, la spécificité est de 99, 5 %.

Ce test réputé « bon » paraît ainsi modérément performant si on reprend la classification subjective présentée précédemment. Deux facteurs majeurs peuvent expliquer les limites de notre étude.

Tout d'abord, au vu de la chronologie des phénomènes virémique, clinique et sérologique que subit un animal nouvellement infecté (figures 1, 2 & 3), l'apparition des anticorps est tardive par rapport au début des symptômes cliniques. En effet, l'incubation de la maladie variant de 2 à 18 jours avec une moyenne de 6 à 7 jours [Lefèvre, 2003], la phase symptomatique qui justifie une déclaration officielle auprès des autorités vétérinaires est bien antérieure à la phase d'apparition des anticorps (figure 3). En règle générale, les tests sérologiques utilisés en vu de l'établissement du diagnostic d'une maladie comprennent deux analyses à 10 - 15 jours d'intervalle afin de mettre en évidence une éventuelle séroconversion. Le manuel de l'OIE recommande par exemple deux analyses à 7 jours d'intervalle minimum. Dans notre cas, on ne peut pas évaluer à quel instant, par rapport à la cinétique des anticorps, le prélèvement a eu lieu. Un défaut de sensibilité était ainsi attendu sans que les qualités intrinsèques du test n'en soient la cause.

A l'inverse, un défaut de spécificité peut être lié à la présence dans notre échantillon de jeunes agnelles non vaccinées et donc incluses dans les analyses mais qui peuvent avoir conservé des anticorps maternels. Au seuil de 53 %, sur les 7 animaux classés en tant que faux-positifs, 5 sont des agnelles provenant de troupeaux vaccinés (adultes). Dans un autre cas, on ne dispose pas d'informations sur l'âge de l'animal. Peu de données sur la durée de persistance des anticorps maternels chez les ovins existent. Une étude de Hong LI *et al.* (1995) met en évidence des anticorps maternels par ELISA de compétition chez des agnelles de 10 semaines +/- 3 semaines. L'équipe d'Abu Elzein (1998) détecte des anticorps maternels qui disparaissent en 4 à 6 mois.

Ces résultats ont conduit à élargir notre échantillon avec des données collectées en zone indemne, c'est-à-dire dont le statut des animaux vis-à-vis de l'infection par le virus était connu : animaux non malades, non infectés. On atteint alors au seuil de 53 % des valeurs de spécificité supérieures à 98 %. On confirme par là que la spécificité du test cELISA II est certainement sous évaluée. On ne peut cependant pas réévaluer la valeur-seuil du test à partir de ces données compte tenu du déséquilibre induit entre les deux types de population : infectés / non infectés.

L'analyse ROC apparaît comme une démarche très pertinente mais nécessiterait d'être appliquée sur des données plus fiables. Le fait de travailler à partir de sérums « tout-venants », sélectionnés sur des critères de disponibilité des informations induit un certain nombre de biais difficiles à évaluer. Notre échantillon notamment ne respecte pas les critères de normalité des populations infectées et non infectées. Bien que l'analyse ROC est reconnue robuste à cet égard [Greiner & *al.*, 2000], notre démarche gagnera à être poursuivie sur des données expérimentales où le statut des animaux vis-à-vis de l'infection est connu et même daté.



De plus, en pratique, le test sert également à des enquêtes de séroprévalence. Il convient alors de disposer de données collectées dans les conditions se rapprochant de celles de ces enquêtes transversales. Les valeurs de sensibilité et de spécificité pourront alors être extrapolées au terrain, mais également avec des précautions.

Le choix de la valeur-seuil devra alors se faire à la lumière de la situation épidémiologique de l'infection dans la zone étudiée. En zone indemne, l'accent sera porté sur la spécificité afin d'obtenir une très bonne valeur prédictive positive (VPP). En effet, les conséquences d'un faux-positif seraient très importantes. Or, compte tenu de la prévalence proche de zéro, seule la spécificité permet d'augmenter la VPP du test. A l'inverse, en zone infectée, on cherche avant tout à localiser tous les malades. La sensibilité du test est alors très importante pour avoir une valeur prédictive négative satisfaisante.

L'utilisation de la séroneutralisation comme test de référence peut également être envisagée. En effet, du fait des campagnes de vaccination successives mises en place depuis 2001 en Corse, la séroprévalence y est très forte. Elle est évaluée à plus de 90 %. Le statut de l'animal à tester n'est donc plus infecté / non infecté mais protégé / non protégé. Dans ce cas, seul le choix d'un test de référence sérologique est possible. On pourra ainsi définir les valeurs de sensibilité et de spécificité, en supposant le test de séroneutralisation comme un test parfait, du test cELISA II dans le cadre d'enquêtes de séroprévalence. On recherchera alors un test avec une très bonne sensibilité afin d'avoir une valeur prédictive négative excellente.

La concordance entre les résultats issus d'un test de séroneutralisation et d'un test cELISA peut également être analysée par le biais d'un test kappa.

Le statut de l'animal peut également être définie en combinant plusieurs tests en série ou en parallèle. Avec les tests en série, on améliore la sensibilité du résultat puisque un animal est déclaré positif si un des deux tests se révèle positif. L'utilisation de tests en parallèle permet d'obtenir une meilleure spécificité. Seuls les animaux obtenant un résultat positif avec le premier test sont recontrôlés avec le second test.

Les autres résultats de notre étude concernent le test cELISA I utilisé lors de l'épizootie de 2000. L'analyse ROC conduite présente les mêmes limites que pour le test cELISA II à ceci près qu'aucun animal n'était vacciné. De plus, on ne dispose des données sérologiques et virologiques que pour les seuls ovins. Les meilleures valeurs de sensibilité et de spécificité sont obtenues pour un seuil de 70 %. Ce résultat varie de manière significative du seuil de 50 % d'inhibition fixé par le fabricant. La sensibilité et la spécificité associées sont respectivement de 84,6 % (IC à 95 % : [72,6 % - 97,1 %]) et 83,9 % (IC à 95 % : [70,9 - 96,8%]).

La comparaison des deux tests conduit à des résultats qu'il est difficile d'interpréter en utilisant le coefficient kappa. En effet, ce coefficient étant très sensible au déséquilibre entre les effectifs dans les différentes classes de résultats, sa valeur ne permet pas de conclure. En revanche, le pourcentage de résultats discordants entre les deux tests donne une information intéressante. En effet, en conservant le seuil à 50 % pour le test cELISA I, ce pourcentage est de 71,2 %. Si ce seuil est redéfini à 70 %, cette discordance devient inférieure à 4 %. Dans l'objectif de se baser sur un test statistique pour comparer ces deux tests et ainsi de s'affranchir de la part de hasard, il est nécessaire d'avoir un meilleur équilibre entre les effectifs des 4 catégories de résultats. Il faut notamment augmenter la part de résultats positifs obtenus avec le test cELISA II. Des recontrôles de sérums de bovins positifs prélevés en 2000 sont envisagés.

La question de l'effet de la congélation/décongélation des sérums sur la qualité des sérums reste également en suspens, ne disposant que de peu d'informations sur le sujet. L'effet de l'expérimentateur doit être quant à lui peu différent des différences possibles d'une analyse à l'autre puisque que c'est le même opérateur qui a effectué les analyses.

Même si l'interprétation de ces résultats est difficile, les deux tests semblent notablement différents. Ce constat conforte les inquiétudes quant à la robustesse et à la répétabilité du test cELISA I.



Dans un volet pratique, la prévalence apparente évaluée en Corse pour l'année 2000 a été revue à la lumière de ces résultats. Il ressort que la prévalence ovine conserve globalement la même distribution géographique mais avec des intensités de prévalence revues à la baisse. Elle est, dans la plupart des communes pour lesquelles les données sont suffisantes, inférieures à 10 % à l'exception de Sartène qui a une prévalence observée supérieure à 20 %. Mais, on ne retrouve aucune commune avec des prévalences supérieures à 50 %.

Au vu des limites inhérentes à notre étude, le calcul d'une prévalence réelle n'a pas de sens. Les valeurs de sensibilité et de spécificité nécessitent d'être réévaluées sous de nouvelles conditions. La connaissance des prévalences réelles reste par ailleurs essentielle pour juger de l'immunité naturelle et vaccinale dans le cadre de l'évaluation de la circulation virale. Une couverture immunitaire de 70-80 % est *a priori* nécessaire pour limiter significativement la circulation virale au sein d'une population.



Les évaluations de test par des approches statistiques mais également de modélisation sont nécessaires à l'épidémiologie. Elles doivent notamment servir à renforcer la surveillance épidémiologique. Une meilleure connaissance des prévalences réelles permet de plus de définir des méthodes de contrôle adaptées et de développer des réseaux d'alerte plus rapide, ce qui va dans le sens d'une gestion des maladies plus efficace. La connaissance des valeurs de sensibilité et de spécificité sert également utilement à la définition des tailles d'échantillon nécessaires pour mener des enquêtes de terrain

Les approches ROC et kappa sont ainsi très intéressantes mais leur interprétation peut se trouver limitée par les données disponibles. La qualité et la représentativité des données restent un problème majeur en épidémiologie du fait des contraintes rencontrées sur le terrain et des biais, en particulier de sélection, difficilement contrôlables. Au delà de ce problème qui nécessitera d'être mieux maîtrisée dans des investigations ultérieures, la juste définition d'un « gold standard » revêt une place primordiale dans ce type d'étude. En effet, elle repose sur une bonne classification des malades et non malades. Si l'outil de référence entraîne des erreurs de classification non évaluées ou mal évaluées, les conclusions de l'étude seront inévitablement erronées. Force est de constater que les tests de référence parfaits sont rares. De même, il est rare de connaître les valeurs de sensibilité et de spécificité des tests de référence. Face à ce constat, d'autres méthodes statistiques ou de modélisation se sont développées afin d'essayer d'évaluer les paramètres intrinsèques d'un test dans des conditions définies sur une population donnée en l'absence de « gold standard ». C'est notamment le cas pour le logiciel TAGS développé par une équipe française [Pouillot *et al.*, 2002]. Un certain nombre de contraintes persistent cependant. Ainsi, il n'a pas été possible de conduire ce genre d'approches avec les données disponibles. En effet, du fait du nombre d'inconnues à évaluer, il aurait valu disposé d'une population supplémentaire avec une prévalence connue et différente de zéro.

La détermination des performances des tests sérologiques de la FCM sera poursuivie en utilisant des données complémentaires et en faisant appel à d'autres techniques statistiques.



## Références bibliographiques

1. Abu Elzein E.M., Aitchison H., Al-Afaleq A.I., Al-Bashir A.M., Ibrahim A.O., Housawi F.M. (1998). A study on bluetongue virus infection in Saudi Arabia using sentinel ruminants. *Onderstepoort J. Vet. Res.*, 65, 243-251.
2. Abu Elzein E.M. (1985). Bluetongue in camels : a serological survey of the one-humped camel (*Camelus dromedarius*) in the Sudan. *Rev Elev Med Vet Pays Trop*, 38(4):438-442.
3. Afshar A. (1994). Bluetongue : laboratory diagnosis. *Comp. Immunol. Microbiol. Infect. Dis.*, 17, 221-242.
4. Afshar A., Thomas F.C., Wright P.F., Shapiro J.L., Anderson J. (1989). Comparison of competitive ELISA, indirect ELISA and standard AGID tests for detecting blue-tongue virus antibodies in cattle and sheep. *Vet. Rec.*, 124, 136-141.
5. Afshar A., Thomas F.C., Wright P.F., Shapiro J.L., Shettigara P.T., Anderson J. (1987). Comparison of competitive and indirect enzyme-linked immunosorbent assays for detection of bluetongue virus antibodies in serum and whole blood. *J. Clin. Microbiol.*, 25, 1705-1710.
6. Billinis C., Koumbati M., Spyrou V., Nomikou K., Mangana O., Panagiotidis C.A., Papadopoulos O. (2001). Bluetongue virus diagnosis of clinical cases by a duplex reverse transcription-PCR: a comparison with conventional methods. *J. Virol. Methods*, 98, 77-89.
7. Bonneau K.R., DeMaula C.D., Mullens B.A., MacLachlan N.J. (2002). Duration of viraemia infectious to *Culicoides sonorensis* in bluetongue virus-infected cattle and sheep. *Vet. Microbiol.*, 88, 115-125.
8. Brennan, P., Silman, A. (1992). Statistical methods for assessing observer variability in clinical measures. *BMJ*, 304, 1491-1494.
9. Briand S., Tricaud Sophie, Michel P. Les limites de l'utilisation du coefficient kappa. adresse URL : <http://www.stat.ucl.ac.be/ISarchives/IScolloques/jsbl2002/Kappamichel.pdf>, consultée le 15 mai 2003.
10. Cameron A.R. and Baldock F.C. (1998). A new probability formula for surveys to substantiate freedom from disease. *Prev. Vet. Med.*, 34(1):1-17.
11. Campbell C.H. and Grubman M.J. (1985). Current knowledge on the biochemistry and immunology of bluetongue. *Prog. Vet. Microbiol. Immunol.*, 1, 58-79.
12. Coetzer J.A.W., Thomson G.R., Tustin R.C. (1994). Bluetongue. In: *Infectious diseases of livestock with special reference to southern Africa*, vol. 1, Oxford University, Oxford, 443-459.
13. Enoe C., Andersen S., Sorensen V., Willeberg P. (2001). Estimation of sensitivity, specificity and predictive values of two serologic tests for the detection of antibodies against *Actinobacillus pleuropneumoniae* serotype 2 in the absence of a reference test (gold standard). *Prev. Vet. Med.*, 51, 227-243.
14. Gaydos J.K., Stallknecht D.E., Kavanaugh D., Olson R.J., Fuchs E.R. (2002) Dynamics of maternal antibodies to hemorrhagic disease viruses (Reoviridae: Orbivirus) in white-tailed deer. *J. Wildl. Dis.*, 38, 253-257.
15. Ghalib H.W., Cherrington J.M., Osburn B.I. (1985). Virological, clinical and serological responses of sheep infected with tissue culture adapted bluetongue virus serotypes 10, 11, 13 and 17. *Vet. Microbiol.*, 10, 179-188.
16. Greiner M. and Gardner I. A. (2000). Application of diagnostic tests in veterinary epidemiologic studies. *Prev. Vet. Med.*, 45(1-2): 43-59.

17. Greiner M., Pfeiffer D., Smith R.D. (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev. Vet. Med.*, 45(1-2): 23-41.
18. Greiner M. and Gardner I. A. (2000). Epidemiologic issues in the validation of veterinary diagnostic tests. *Prev. Vet. Med.*, 45(1-2): 3-22.
19. Greiner M. and Gardner I. A. (1999). Roc analysis. In: *Advanced methods for test validation and interpretation in veterinary medicine*. Berlin, 78 pp.
20. Greiner M. (1996). Two-graph receiver operating characteristic (TG-ROC): update version supports optimisation of cut-off values that minimise overall misclassification costs. *J. Immunol. Methods*, 191, 93-94.
21. Greiner M., Sohr D., Gobel P. (1995). A modified ROC analysis for the selection of cut-off values and the definition of intermediate results of serodiagnostic tests. *J. Immunol. Methods*, 185, 123-132.
22. Greiner M. (1995). Two-graph receiver operating characteristic (TG-ROC): a Microsoft-EXCEL template for the selection of cut-off values in diagnostic tests. *J. Immunol. Methods*, 185, 145-146.
23. Hamblin C., Salt J.S., Graham S.D., Hopwood K., Wade-Evans A.M. (1998). Bluetongue virus serotypes 1 and 3 infection in Poll Dorset sheep. *Aust. Vet. J.*, 76, 622-629.
24. Li H., Shen D.T., O'Toole D., Knowles D.P., Gorham J.R., Crawford T.B. (1995). Investigation of sheep-associated malignant catarrhal fever virus infection in ruminants by PCR and competitive inhibition enzyme-linked immunosorbent assay. *J. Clin. Microbiol.*, 33(8):2048-2053.
25. Katz J., Alstad D., Gustafson G., Evermann J. (1994). Diagnostic analysis of the prolonged bluetongue virus RNA presence found in the blood of naturally infected cattle and experimentally infected sheep. *J. Vet. Diagn. Invest.*, 6, 139-142.
26. Katz J.B., Gustafson G.A., Alstad A.D., Adler K.A., Moser, K.M. (1993). Colorimetric diagnosis of prolonged bluetongue viraemia in sheep, using an enzyme-linked oligonucleotide sorbent assay of amplified viral nucleic acids. *Am. J. Vet. Res.*, 54, 2021-2026.
27. Koumbati M., Mangana O., Nomikou K., Mellor P.S., Papadopoulos O. (1999). Duration of bluetongue viraemia and serological responses in experimentally infected European breeds of sheep and goat. *Vet. Microbiol.*, 64, 277-285.
28. Le Fur C. et Eddeiraï D. (2002). Enquête sérologique post-vaccinale sur les troupeaux transhumants en Corse en 2001. Mémoire de CES d'épidémiologie animale, 19 pp.
29. Lefèvre P.-C. (2003). Fièvre catarrhale du mouton. In: *Principales maladies infectieuses et parasitaires du bétail, Europe et régions chaudes*. Tec & Doc, M.I. (Ed.), Londres, Paris, New York, 667-686.
30. Luedke A.J. (1969). Bluetongue in sheep: viral assay and viraemia. *Am. J. Vet. Res.*, 30, 499-509.
31. MacLachlan N.J., Nunamaker R.A., Katz J.B., Sawyer M.M., Akita G.Y., Osburn B.I. & Taberchnick W.J. (1994). Detection of bluetongue virus in the blood of inoculated calves: comparison of virus isolation, PCR assay and *in vitro* feeding of *Culicoides variipennis*. *Arch. Virol.*, 136, 1-8.
32. Meiswinkel R., Nevill E.M., Venter G.J. (1994). Vectors : *Culicoides spp.* In : *Infectious diseases of livestock with special reference to southern Africa*. Coetzer J.A.W., Thomson G.R. & Tustin R.C. (Ed.), Oxford University Press, Capetown (Afrique du Sud), 68-89
33. Mellor P.S. and Pitzolis, G. (1979). Observations on breeding sites and light-trap collections of *Culicoides* during an outbreak of bluetongue in Cyprus. *Bull. Ent. Res.*, 69, 229-234.

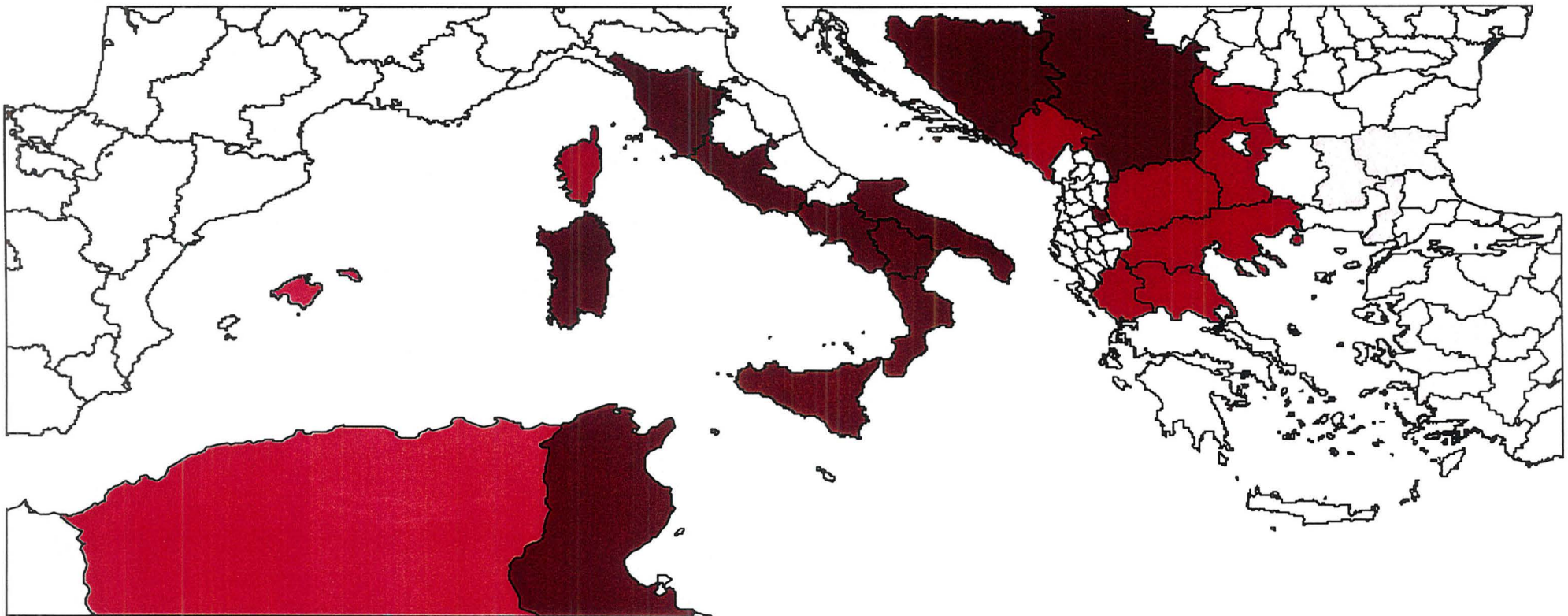


34. Office International des Épizooties (2002). Principles of validation of diagnostics assays for infectious diseases. *In: Manual of standards for diagnostics tests and vaccines*, OIE (Ed.), Paris, 8-15.
35. Pouillot R., Gerbier G., Gardner I. A. (2002). TAGS, a program for the evaluation of test accuracy in the absence of a gold standard. *Prev. Vet. Med.*, 53, 67-81.
36. Rogan W.J., Gladen B. (1978). Estimating prevalence from the results of a screening test. *Am. J. Epidemiol.*, 107 (1), 71-76.
37. Roger F., Tatem A., de la Rocque S., Hendrihx P., Baylis M., Delécolle J.-P., Rogers D. (2002). Modelling locations at risks for potential vectors of Bluetongue using satellite imagery. Poster, non publié.
38. Simpson V.R. (1979). Bluetongue antibody in Botswana's domestic and game animals. *Trop. Anim. Health Prod.*, 11(1): 43-49.
39. Toma B., Dufour B., Sanaa M., Bénét J.J., Ellis P., Moutou F., Louza A. (1996). Le dépistage des maladies infectieuses animales. *In : Épidémiologie appliquée à la lutte collective contre les maladies animales transmissibles majeures*. AEEMA (Ed.), Maisons-Alfort, 41-86.
40. Thrusfield M, Ortega C, de Blas I, Noordhuizen JP, Frankena K. (2001). WIN EPISCOPE 2.0 : improved epidemiological software for veterinary. *Vet. Rec.*, 148(18): 567-72.
41. Thrusfield M. (1995). *Veterinary epidemiology*. 2<sup>nd</sup> edition, Blackwell Science, Oxford (England), 479 pp.
42. Uren M.F., Squire K.R. (1982). The clinico-pathological effect of bluetongue virus serotype 20 in sheep. *Aust. Vet. J.*, 58, 11-15.
43. Verwoerd D.W. and Erasmus B.J. (1994). Bluetongue. *In: Infectious disease do livestock with special reference to southern Africa*. Coetzer J.A.W., Thomson G.R. & Tustin .R.C. (Ed.), Oxford University Press, Capetown (Afrique du Sud), 443-459.
44. Xu H., Lohr J., Greiner M. (1997). The selection of ELISA cut-off points for testing antibody to Newcastle disease by two-graph receiver operating characteristic (TG-ROC) analysis. *J. Immunol. Methods*, 208, 61-64.
45. Zientara S., Sailleau C., Dauphin G., Roquier C., Remond E.M., Lebreton F., Hammoumi S., Dubois, E., Agier C., Merle G., Breard E. (2002). Identification of bluetongue virus serotype 2 (Corsican strain) by reverse-transcriptase PCR reaction analysis of segment 2 of the genome. *Vet. Rec.* 150, 598-601.
46. Zientara S., Gourreau J.-M. (2001). La fièvre catarrhale du mouton. *Virologie*, 5, 449-451.
47. Zientara S., Grillet C., de la Rocque S., Gourreau J.M., Grégory M., Hendrickx P., Libeau P., Sailleau C., Albina E., Bréard E. and Delécolle J.C. (2001). La fièvre catarrhale ovine en Corse en 2001. *Épidémiol. et Santé anim.*, 40 : 129-137.
48. Zientara S., de La Rocque S., Gourreau J.M., Grégory M., Diallo A., Hendrickx P., Libeau G., Sailleau C., Delécolle J.C. (2000). La fièvre catarrhale ovine en Corse en 2000. *Épidémiol. et santé anim.* 38, 133-144.

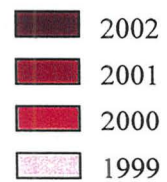
# ANNEXES



## ANNEXE I



Carte 1 : Localisation des derniers foyers recensés de fièvre catarrhale du mouton dans le bassin méditerranéen (données 2002) – source OIE



### Manual of standards Diagnostic Tests and Vaccines 2000

#### CHAPTER I.1.3.

### PRINCIPLES OF VALIDATION OF DIAGNOSTIC ASSAYS FOR INFECTIOUS DISEASES

---

#### INTRODUCTION

*Validation is the evaluation of a process to determine its fitness for a particular use. A validated assay yields test results that identify the presence of a particular analyte (e.g. an antibody) and allows predictions to be made about the status of the test subject. However, for infectious disease diagnostic assays, the identity and definition of the criteria required for assay validation are elusive, and the process leading to a validated assay is not standardised.*

*By considering the variables that affect an assay's performance, the criteria that must be addressed in assay validation become clearer. The variables can be grouped into three categories: (a) the sample - host/organism interactions affecting the analyte composition and concentration in the serum sample; (b) the assay system - physical, chemical, biological and technician-related factors affecting the capacity of the assay to detect a specific analyte in the sample; and (c) the test result - the capacity of a test result, derived from the assay system, to predict accurately the status of the host relative to the analyte in question.*

*Factors that affect the concentration and composition of analyte in the serum sample are mainly attributable to the host and are either inherent (e.g. age, sex, breed, nutritional status, pregnancy, immunological responsiveness) or acquired (e.g. passively acquired antibody, active immunity elicited by vaccination or infection). Nonhost factors, such as contamination or deterioration of the sample, may also affect the analyte in the sample.*

*Factors that interfere with the analytical accuracy of the assay system include instrumentation, technician error, reagent choice (both chemical and biological) and calibration, accuracy and acceptance limits of controls, reaction vessels, water quality, pH and ionicity of buffers and diluents, incubation temperatures and durations, and error introduced by detection of closely related analytes, such as antibody to cross-reactive organisms, rheumatoid factor, or heterophile antibody.*

*Factors that influence the capacity of the test result to predict accurately the infection or analyte status of the host are diagnostic sensitivity (D-SN), diagnostic specificity (D-SP), and prevalence of the disease in the population targeted by the assay. (In this chapter, the terms 'positive' and 'negative' have been reserved for test results and never refer to infection or antibody/antigen status of the host. Whenever reference is made to 'infection' or 'analyte', any method of exposure to an infectious agent that could be detected directly [e.g. antigen] or indirectly [e.g. antibody] by an assay, should be inferred.) D-SN and D-SP are derived from test results on samples obtained from selected reference animals. The methods used to select the reference animals are critical to the accuracy of the estimates (4). The degree to which the reference animals represent all of the host and environmental variables in the population targeted by the assay has a major impact on the accuracy of test-result interpretation. For example, experienced diagnosticians are aware that an assay, validated by using samples from northern European cattle, may not give valid results for the distinctive populations of cattle in Africa.*

*The capacity of a positive or negative test result to predict accurately the infection status of the animal is the most important consideration of assay validation. This capacity is not only dependent on a highly precise and accurate assay and carefully derived estimates of D-SN and D-SP, but is heavily influenced by prevalence of the infection in the targeted population or the likelihood that an animal is infected based on clinical criteria. Without a current estimate of the disease prevalence in that population or likelihood of infection in an individual animal, the interpretation of a positive or negative test result may be compromised.*



*Many variables obviously must be addressed before an assay can be considered to be 'validated' (4, 10). However, there is no consensus whether the concept of assay validation is a time-limited process during which only those factors intrinsic to the assay are optimised and standardised, or whether the concept includes an ongoing validation of assay performance for as long as the assay is used. Accordingly, the term 'validated assay' elicits various interpretations among laboratory diagnosticians and veterinary clinicians. Therefore, a working definition of assay validation is offered as a context for the guidelines outlined below. Ideally, all diagnostic assays would be fully validated, but in practice there are sometimes limitations to the completeness of validation.*

## A. DEFINITION OF ASSAY VALIDATION

A validated assay consistently provides test results that identify animals as positive or negative for an analyte or process (e.g. antibody, antigen, or induration at skin test site) and, by inference, accurately predicts the infection status of animals with a predetermined degree of statistical certainty. This chapter will focus on the principles underlying development and maintenance of a validated assay. Guidelines for the initial stages in assay development are included because they constitute part of the validation process. How this early process is carried out heavily influences the capacity of the eventual test result to provide diagnostic accuracy.

It must be emphasised that an assay, when applied to target populations, will minimise misclassifications of animals as false positive or false negative only to the extent that validity is assured for all stages of the assay validation process. This assumes that a well designed and documented test method and proper reagents, in combination with well trained technicians, will give a stable assay within the laboratory. It also assumes a thorough use of the most rigorous experimental design and epidemiological and statistical tools. These are required to reduce bias, random error, and false assumptions about the reference population of animals upon which the assay performance estimates are made (4).

The working definition of assay validation used in this paper generalises a process that takes on many different forms depending on the intended use of the assay being developed. No matter what the form of the assay, without scientific rigour in the validation process, the assay will not perform to expectations. However, there are many traditional techniques in widespread use that have not been subjected to the full validation process.

## B. STAGES OF ASSAY VALIDATION

Development and validation of an assay is an incremental process consisting of at least five stages:

- 1) Determination of the feasibility of the method for a particular use;
- 2) Choice, optimisation, and standardisation of reagents, techniques and methods;
- 3) Determination of the assay's performance characteristics;
- 4) Continuous monitoring of assay performance; and
- 5) Maintenance and enhancement of validation criteria during routine use of the assay (Figure 1)

Although some scientists may question the relevance of the fourth and fifth stages as validation criteria, they are included here because an assay can be considered to be valid only to the extent that test results are valid, i.e. whether they fall within statistically defined limits and provide accurate inferences. In practice, there are many traditional techniques in widespread use that have not been through the formal validation process given above. In such cases the process may appropriately be entered at stages 4 and 5, thus building up a data set of validation data for assays that lack the necessary information for stages 1-3.

An indirect enzyme-linked immunosorbent assay (ELISA) test for detection of antibody will be used in this chapter to illustrate the principles of assay validation. It is a test format that can be difficult to validate because of signal amplification of both specific and nonspecific components (2). This methodology serves to highlight the problems that need to be addressed in any assay validation process. The same basic principles are used in validation of other complex or simple assay formats.

Because of space limitations, this introductory chapter provides only basic guidelines for the principles concerned with assay validation. It is derived from a more detailed treatment of the subject (7).



## STAGE 1. FEASIBILITY STUDIES

In the ELISA example, feasibility studies are the first stage in validating a new assay. They are carried out in order to determine if the selected reagents and protocol have the capacity to distinguish between a range of antibody concentrations to an infectious agent while providing minimal background activity. Feasibility studies also give initial estimates of repeatability, and of analytical sensitivity and specificity.

### 1. Control samples

It is useful to select four or five samples (serum in our example) that range from high to low levels of antibodies against the infectious agent in question. In addition, a sample containing no antibody is required. These samples will be used to optimise the assay reagents and protocol during feasibility studies, and later as control samples. The samples ideally should represent known infected and uninfected animals from the population that eventually will become the target of the validated assay. The samples should have given expected results in one or more serological assay(s) other than the one being validated. The samples are preferably derived from individual animals, but they may represent pools of samples from several animals. A good practice is to prepare a large volume (e.g. 10 ml) of each sample and divide it into 0.1 ml aliquots for storage at or below  $-20^{\circ}\text{C}$ . One aliquot of each sample is thawed, used for experiments, and ideally then discarded. If it is impractical to discard the aliquot, it may be held at  $4^{\circ}\text{C}$  between experiments for up to approximately 2 weeks; however, there is a possibility of sample deterioration under these circumstances. Then, another is thawed for further experimentation. This method provides the same source of serum with the same number of freeze/thaw cycles for all experiments (repeated freezing and thawing of serum can denature antibodies so should be avoided). Also, variation is reduced when the experimenter uses the same source of serum for all experiments rather than switching among various sera between experiments. This approach has the added advantage of generating a data trail for the repeatedly run samples. After the initial stages of assay validation are completed, one or more of the samples can become the serum control(s) that are the basis for data expression and repeatability assessments both within and between runs of the assay. They may also serve as standards if their activity has been predetermined; such standards provide assurance that runs of the assay are producing accurate data (10).

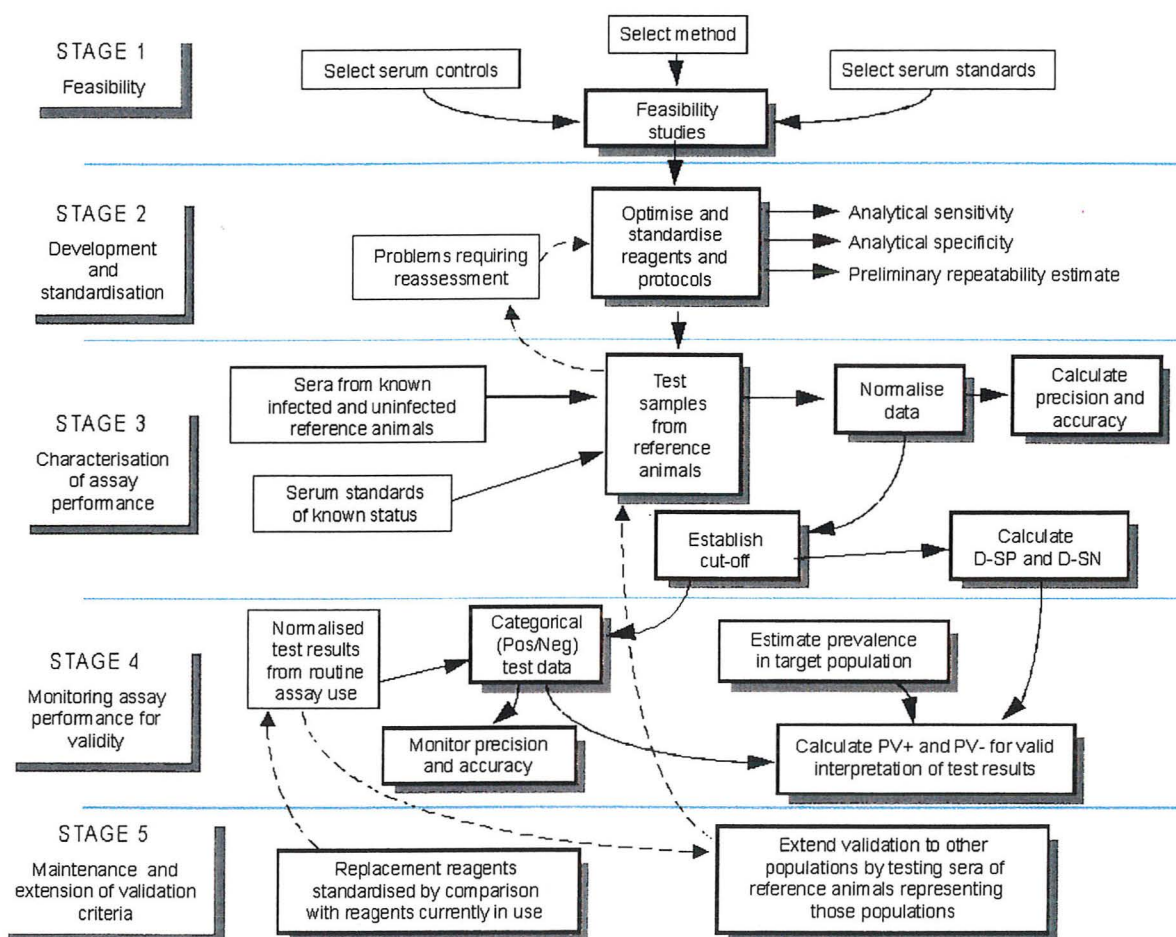
It is highly desirable to include international standard sera if they are available. This may lead to harmonisation between the assay under development and a standard test method in which international standard sera are normally used (9).

### 2. Selection of method to achieve normalised results

Normalisation adjusts raw test results of all samples relative to values of controls included in each run of the assay (not to be confused with transformation of data to achieve a 'normal' [Gaussian] distribution). The method of normalisation and expression of data should be determined, preferably no later than at the end of the feasibility studies. Comparisons of results from day to day and between laboratories are most accurate when normalised data are used. For example, in ELISA systems, raw optical density (absorbance) values are absolute measurements that are influenced by ambient temperatures, test parameters, and photometric instrumentation. To account for this variability, results are expressed as a function of the reactivity of one or more serum control samples that are included in each run of the assay. Data normalisation is accomplished in the indirect ELISA by expressing absorbance values in one of several ways (10). A simple and useful method is to express all absorbance values as a percentage of a single high-positive serum control that is included on each plate. This method is adequate for most applications. More rigour can be brought to the normalisation procedure by calculating results from a standard curve generated by several serum controls. It requires a more sophisticated algorithm, such as linear regression or log-logit analysis. This approach is more precise because it does not rely on only one high-positive control sample for data normalisation, but rather uses several serum controls, adjusted to expected values, to plot a standard curve from which the sample value is extrapolated. This method also allows for exclusion of a control value that may fall outside expected confidence limits.



For assays that are end-pointed by sample titration, such as serum (viral) neutralisation, each run of the assay is accepted or rejected based on whether control values fall within predetermined limits. Because sample values usually are not adjusted to a control value, the data are not normalised by the strict definition of the term.



**Fig. 1.** Stages in the incremental process of assay validation.

Whatever method is used for normalisation of the data, it is essential to include additional controls for any reagent that may introduce variability and thus undermine attempts to achieve a validated assay. The normalised values for those controls need to fall within predetermined limits (e.g. within an appropriate multiple of the standard deviation of the mean of many runs of each control). The chosen limits should reflect a reasonable and tolerable assay run rejection rate and an acceptable risk that some test samples may be misclassified.

## STAGE 2. ASSAY DEVELOPMENT AND STANDARDISATION

When the feasibility of the method has been established, the next step is to proceed with development of the assay and standardisation of the selected reagents and protocols.

### 1. Selection of optimal reagent concentrations and protocol parameters

Optimal concentrations/dilutions of the antigen adsorbed to the plate, serum, enzyme/antibody conjugate, and substrate solution are determined through 'checkerboard' titrations of each reagent against all other reagents, following confirmation of the best choice of reaction vessels (usually evaluation of two or three types of microtitre plates, each with its different binding characteristics, to minimise background activity while achieving the maximum spread in activity between negative and high-positive samples). Additional experiments determine the optimal temporal, chemical, and physical variables in the protocol, including incubation temperatures and durations; the type, pH, and molarity of diluent, washing and blocking buffers; and equipment used in each step of the assay (for instance pipettes and washers that give the best reproducibility).



Optimisation of the reagents and protocol includes an assessment of accuracy by incorporating, in each run of the assay, one or more serum standards of a known level of activity for the analyte in question. An optimised assay that repeatedly achieves the same results for a serum standard and the serum controls may be designated as a standardised assay.

## **2. Repeatability - preliminary estimates**

Preliminary evidence of repeatability (agreement between replicates within and between runs of the assay) is necessary to warrant further development of the assay. This is accomplished by evaluating results from replicates of all samples in each plate (intraplate variation), and interplate variation using the same samples run in different plates within a run and between runs of the assay. For ELISA, raw absorbance values are usually used at this stage of validation because it is uncertain whether the results of the high-positive control serum, which could be used for calculating normalised values, are reproducible in early runs of the assay format. Also, expected values for the controls have not yet been established. Three-to-four replicates of each sample run in at least five plates on five separate occasions are sufficient to provide preliminary estimates of repeatability. Coefficients of variation (standard deviation of replicates/mean of replicates), generally less than 20% for raw absorbance values, indicates adequate repeatability at this stage of assay development. However, if evidence of excessive variation (>30%) is apparent for most samples within and/or between runs of the assay, more preliminary studies should be done to determine whether stabilisation of the assay is possible, or whether the test format should be abandoned. This is important because an assay that is inherently variable has a high probability of not withstanding the rigours of day-to-day testing on samples from the targeted population of animals.

## **3. Determination of analytical sensitivity and specificity**

The analytical sensitivity of the assay is the smallest detectable amount of the analyte in question, and analytical specificity is the degree to which the assay does not cross-react with other analytes. These parameters are distinguished from diagnostic sensitivity and specificity as defined below. Analytical sensitivity can be assessed by end-point dilution analysis, which indicates the dilution of serum in which antibody is no longer detectable. Analytical specificity is assessed by use of a panel of sera derived from animals that have experienced related infections that may stimulate cross-reactive antibodies. If, for instance, the assay does not detect antibody in limiting dilutions of serum with the same efficiency as other assays, or cross-reactivity is common when sera from animals with closely related infections are tested, the reagents need to be recalibrated or replaced, or the assay should be abandoned. However, there is a place for a test with low specificity, providing the sensitivity is high, for use as a 'screening test' in situations where a large number of samples have to be screened and where there is a 'confirmatory test' that has high specificity.

### **STAGE 3. DETERMINING ASSAY PERFORMANCE CHARACTERISTICS**

If the feasibility and initial development and standardisation studies indicate that the assay has potential for field application, the next step is to identify the assay's performance characteristics.

#### **1. Diagnostic sensitivity and specificity**

##### ***a) Principles and definitions***

Estimates of diagnostic sensitivity (D-SN) and diagnostic specificity (D-SP) are the primary parameters obtained during validation of an assay. They are the basis for calculation of other parameters from which inferences are made about test results. Therefore, it is imperative that estimates of D-SN and D-SP be as accurate as possible. Ideally, they are derived from testing a series of samples from reference animals of known history and infection status relative to the disease/infection in question. Diagnostic sensitivity is the proportion of known infected reference animals that test positive in the assay; infected animals that test negative are considered to have false-negative results. Diagnostic specificity is the proportion of uninfected reference animals that test negative in the assay; uninfected reference animals that test positive are considered to have false-positive results. The number and source of reference samples used to derive D-SN and D-SP are of paramount importance if the assay is ever to be properly validated for use in the general population of animals targeted by the assay.



It is possible to calculate the number of reference samples, from animals of known infection/exposure status, required for determinations of D-SN and D-SP that will have statistically defined limits. Formulae for these calculations, their limitations, and a discussion of the selection criteria for standard sera are detailed elsewhere (4, 7). Because of the many variables that must be accounted for, it is desirable to include at least 300 reference samples from known-infected animals, and about 1000 samples from known-uninfected animals should be included to determine initial estimates of D-SN and D-SP, respectively. Such a large number of reference animals may be difficult to achieve. Starting with fewer animals followed by updating of D-SN and D-SP estimates as more reference animals become available may be the only practical way to get initial estimates of D-SN and D-SP (See Stage 5 below).

*b) Standards of comparison for the new assay*

In serology, the 'standard of comparison' is the results of a method or combination of methods with which the new assay is compared. Although the term 'gold standard' is commonly used to describe any standard of comparison, it should be limited to methods that unequivocally classify animals as infected or uninfected. Some isolation methods themselves have problems of repeatability and sensitivity. Gold standard methods include unequivocal isolation of the agent or pathognomonic histopathological criteria. Because a true gold standard is impossible to achieve, relative standards of comparison are often necessary; these include results from other serological assays and from experimentally infected or vaccinated animals. Calculations of D-SN and D-SP are most reliable when the gold standard of comparison is available. When only relative standards of comparison are available, estimates of D-SN and D-SP for the new assay may be compromised because the error in the estimates of D-SN and D-SP for the relative standard is carried over into those estimates for the new assay.

*c) Precision, repeatability, reproducibility, and accuracy*

Repeatability and reproducibility are estimates of precision in the assay. Precision is a measure of dispersion of results for a repeatedly tested sample; a precise assay has only a small amount of dispersion. Repeatability in a diagnostic assay has two elements: the amount of agreement between replicates (usually two or three) of each sample within a run of the assay, and the amount of between-run agreement for the normalised values of each control sample. Reproducibility is the amount of agreement between results of the same samples tested in different laboratories. Accuracy is the amount of agreement between a test value and the expected value for an analyte in a standard sample of known activity (e.g. titre or concentration). An assay system may be precise, but not accurate, if the test results do not agree with the expected value for the standard.

Reliable estimates of repeatability and accuracy, both within and between runs of the assay, can be obtained by use of normalised results from the many runs of the new assay that were required to assess the sera of reference animals (less reliable estimates were obtained from preliminary data using raw absorbance values). At least 10, and preferably 20 runs of the assay will give reasonable initial estimates of these parameters. Methods for evaluating these parameters have been described in detail (7).

**Accuracy** can be assessed by inclusion of one or more standards (samples of known titre, concentration, etc.) in each run of the assay. The standards may be control sera provided that the amount of analyte (e.g. titre, concentration) in each one has been previously determined by comparison with primary or secondary reference standards (10), and the control sera are not used in the data normalisation process.

**Reproducibility** of the assay is determined in several laboratories using the identical assay (protocol, reagents, and controls) on a group of at least 10 samples, preferably duplicated to a total of 20 samples. These samples need to represent the full range of expected analyte concentrations in samples from the target population. The extent to which the collective results for each sample deviate from expected values is a measure of assay reproducibility. The degree of concordance of between-laboratory data is one more basis for determining whether the assay's performance characteristics are adequate to constitute a validated assay.



## 2. Selection of the cut-off (positive/negative threshold)

To achieve estimates of D-SN and D-SP of the new assay, the test results first must be reduced to positive or negative categories. This is accomplished by insertion of a cut-off point (threshold or decision limit) on the continuous scale of test results. Although many methods have been described for this purpose, three examples will illustrate different approaches, together with their advantages and disadvantages. The first is a cut-off based on the frequency distributions (7) of test results from uninfected and infected reference animals. This cut-off can be established by visual inspection of the frequency distributions, by receiver-operator characteristics (ROC) analysis (5, 11), or by selection that favours either D-SN or D-SP, whichever is required for a given assay (8). A second approach is establishing a cut-off based only on uninfected reference animals; this provides an estimate of D-SP but not D-SN. The third method provides an 'intrinsic cut-off' based on test results from sera drawn randomly from within the target population with no prior knowledge of the animals' infection status (3). Although no estimates of D-SN and D-SP are obtained by this method, they can be determined as confirmatory data are accumulated.

If considerable overlap occurs in the distributions of test values from known infected and uninfected animals, it is difficult to select a cut-off that will accurately classify these animals according to their infection status. Rather than a single cut-off, two cut-offs can be selected that define a high D-SN (e.g. inclusion of 99% of the values from infected animals), and a high D-SP (e.g. 99% of the values from uninfected animals). The values that fall between these percentiles would then be classified as suspicious or equivocal, and would require testing by a confirmatory assay or retesting for detection of seroconversion.

## 3. Calculation of diagnostic sensitivity and specificity

The selection of a cut-off allows classification of test results into positive or negative categories. Calculation of D-SN and D-SP are aided by associating the positive/negative categorical data with the known infection status for each animal using a two-way (2 x 2) table (Table 1). After the cut-off is established, results of tests on standard sera can be classified as true positive (TP) or true negative (TN) if they are in agreement with those of the gold standard (or other standard of comparison). Alternatively, they are classified as false positive (FP) or false negative (FN) if they disagree with the standard. Diagnostic sensitivity is calculated as  $TP/(TP + FN)$  whereas diagnostic specificity is  $TN/(TN + FP)$ ; the results of both calculations are usually expressed as percentages (Table 1).

**Table 1.** Calculations of D-SN and D-SP aided by a 2 x 2 table that associates infection status with test results from 2000 reference animals

		Reference animals of known infection status	
		Infected (n = 600)	Uninfected (n = 1400)
Test Result	POSITIVE	570	46
	Negative	30	1354
		TP	FP
		FN	TN
Diagnostic sensitivity		Diagnostic specificity	
$\frac{TP}{TP + FN} = \frac{570}{600} = 95.0\%$		$\frac{TN}{TN + FP} = \frac{1354}{1400} = 96.7\%$	



#### **4. Harmonisation of assays**

When an international standard method (9) is available for detection of an analyte, it is possible to compare the performance of that method with the one under development. This process requires use of the same serum controls and/or standards in both assays. If international standard sera are available (preferably at least three: negative, weak positive and strong positive), they should be included in the assay-comparison study. This could lead to a new assay that is indexed to an international standard method and international standard sera (9). Harmonisation of the two assays may then be realised.

#### **STAGE 4. MONITORING VALIDITY OF ASSAY PERFORMANCE**

##### **Interpretation of test results - factors affecting assay validity**

An assay's test results are useful only if the inferences made from them are accurate. A common error is to assume that an assay with 99% D-SN and 99% D-SP will generate one false-positive and one false-negative result for approximately every 100 tests on animals from the target population. Such an assay may be precise and accurate, but produce test results that do not accurately predict infection status. For example, if the prevalence of disease in a population targeted by the assay is only 1 per 1000 animals, and the false-positive test rate is 1 per 100 animals (99% D-SP), for every 1000 tests on that population, ten will be false positive and one will be true positive. Hence, only approximately 9% of positive test results will accurately predict the infection status of the animal; the positive test results will misclassify the animal 91% of the time. This illustrates that the capacity of a positive or negative test result to predict infection status is dependent on the prevalence of the infection in the target population (6). Of course, the prevalence will probably have been determined by use of a serological test.

An estimate of prevalence in the target population is necessary for calculation of the predictive values of positive (PV+) or negative (PV-) test results. When test values are reported without providing estimates of the assay's D-SP and D-SN, it is not possible to make informed predictions of infection status from test results (6). It is, therefore, highly desirable to provide an interpretation statement with test results accompanied by a small table indicating PV+ and PV- for a range of expected prevalences of infection in the target population. Without provision of such information, test results from the assay may have failed to classify accurately the infection status of animals, and thus do not reflect a fully validated assay.

#### **STAGE 5. MAINTENANCE AND ENHANCEMENT OF VALIDATION CRITERIA**

A validated assay needs constant monitoring and maintenance to retain that designation. Once the assay is put into routine use, internal quality control is accomplished by consistently monitoring the assay for assessment of repeatability and accuracy (1).

Reproducibility between laboratories should be assessed at least twice each year. It is useful to volunteer membership in a consortium of laboratories that are interested in evaluating their output. In the near future, good laboratory practice, including implementation of a total quality assurance programme, will become essential for laboratories seeking to meet national and international certification requirements (see Chapter I.1.2.).

Proficiency testing is a form of external quality control for an assay. It is usually administered by a reference laboratory that distributes panels of samples, receives the results from the laboratories, analyses the data, and reports the results back to the laboratories. If results from an assay at a given laboratory remain within acceptable limits and show evidence of accuracy and reproducibility, the laboratory may be certified by government agencies or reference laboratories as an official laboratory for that assay. Panels of sera for proficiency testing should contain a full representation of an analyte's concentration in animals of the target population. If the panels only have high-positive and low-positive sera (with none near the assay's cut-off), the exercise will only give evidence of reproducibility at the extremes of analyte concentration, and will not clarify whether routine test results on the target population properly classify infection status of animals.



Because of the extraordinary set of variables that impact on the performance of serodiagnostic assays, it is highly desirable to expand the number of standard sera from animals of known infection status because of the principle that error in the estimates of D-SN and D-SP is reduced with increasing sample size. Furthermore, when the assay is to be transferred to a completely different geographical region, it is essential to re-validate the assay by subjecting it to sera from populations of animals that reside under local conditions.

When a serum control sample is nearing depletion, it is essential to prepare and repeatedly test a replacement before the serum control is depleted. The prospective control sample should be included in 10-20 runs of the assay before depletion of the original control to establish its proportional relationship to the nearly depleted control. If the depleted sample was a positive control in ELISAs where the normalised value is expressed as a per cent of that positive control, the proportional difference in ELISA activity between the original and replacement sera must be factored into the normalisation algorithm to retain the same cut-off, and thus the same D-SN and D-SP in the assay. When other reagents, such as antigen for capture of antibody, must be replaced, they should be produced using the same criteria as for the original reagents, and tested in at least five runs of the assay using a panel of sera that has been designed for this purpose. Whenever possible, it is important to change only one reagent at a time to avoid the compound problem of evaluating more than one variable at a time.

### C. VALIDATION OF ASSAYS OTHER THAN ENZYME-LINKED IMMUNOSORBENT ASSAY

Although the example used has been an indirect ELISA test, the same principles apply to the validation of any diagnostic assay. It is of utmost importance not to stop after the first two stages of assay validation - that does not constitute a validated assay for diagnostic use. Although reagent and protocol refinement are important, the selection of the reference populations is probably the most critical factor. It is no surprise when reviewing the literature to find a wide range of estimates for D-SN and D-SP for the same basic assay. Although part of the variation may be attributed to the reagents chosen, it is likely that the variation in estimates of D-SN and D-SP is due to biased selection of sera on which the test was 'validated'. This stage in assay validation needs more attention than it has been given previously. This is particularly true in the current atmosphere of international trade agreements and all their implications with regard to the movement of animals and animal products.

### REFERENCES

1. Cembrowski G.S. & Sullivan A.M. (1992). Quality control and statistics. *In: An Introduction to Clinical Chemistry*, Bishop M.L., Duben-Engelkirk J.L. & Fody E.P., eds. Lippincott, Philadelphia, USA, 63-101.
2. Crowther J.R. (1995). ELISA theory and practice. *In: Methods in Molecular Biology*. Humana Press, Totowa, NJ, USA, 1-256.
3. Greiner M., Franake C.R., Bohning D. & Schlattmann P. (1994). Construction of an intrinsic cut-off value for the sero-epidemiological study of *Trypanosoma evansi* infections in a canine population in Brazil: a new approach towards unbiased estimation of prevalence. *Acta Trop.*, **56**, 97-109.
4. Greiner M. & Gardner I. (2000). Epidemiologic issues in the validation of veterinary diagnostic tests. *Vet. Prev. Med.*, **45**, 3-22.
5. Greiner M., Pfeiffer D. & Smith R.D. (2000). Principles and practical application of the receiver operating characteristic (ROC) analysis for diagnostic tests. *Vet. Prev. Med.*, **45**, 23-41.
6. Jacobson R.H. (1991). How well do serodiagnostic tests predict the infection of disease status of cats. *J. Am. Vet. Med. Assoc.*, **199**, 1343-1347.
7. Jacobson R.H. (1998). Validation of serological assays for diagnosis of infectious diseases. *Rev. sci. tech. Off. int. Epiz.*, **17**, 469-486.
8. Smith R.D. (1991). Clinical Veterinary Epidemiology. Butterworth-Heinemann, Stoneham, MA, USA, 1-223.
9. Wright P.F. (1998). International standards for test methods and reference sera for diagnostic tests for antibody detection. *Rev. sci. tech. Off. int. Epiz.*, **17**, 527-533.
10. Wright P.F., Nilsson E., Van Rooij E.M.A., Lelenta M. & Jeggo M.H. (1993). Standardization and validation of enzyme-linked immunosorbent assay techniques for the detection of antibody in infectious disease diagnosis. *Rev. sci. tech. Off. int. Epiz.*, **12**, 435-450.
11. Zweig M.H. & Campbell G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.*, **39**, 561-577.



## ANNEXE III

### Morphologie du virus de la fièvre catarrhale du mouton

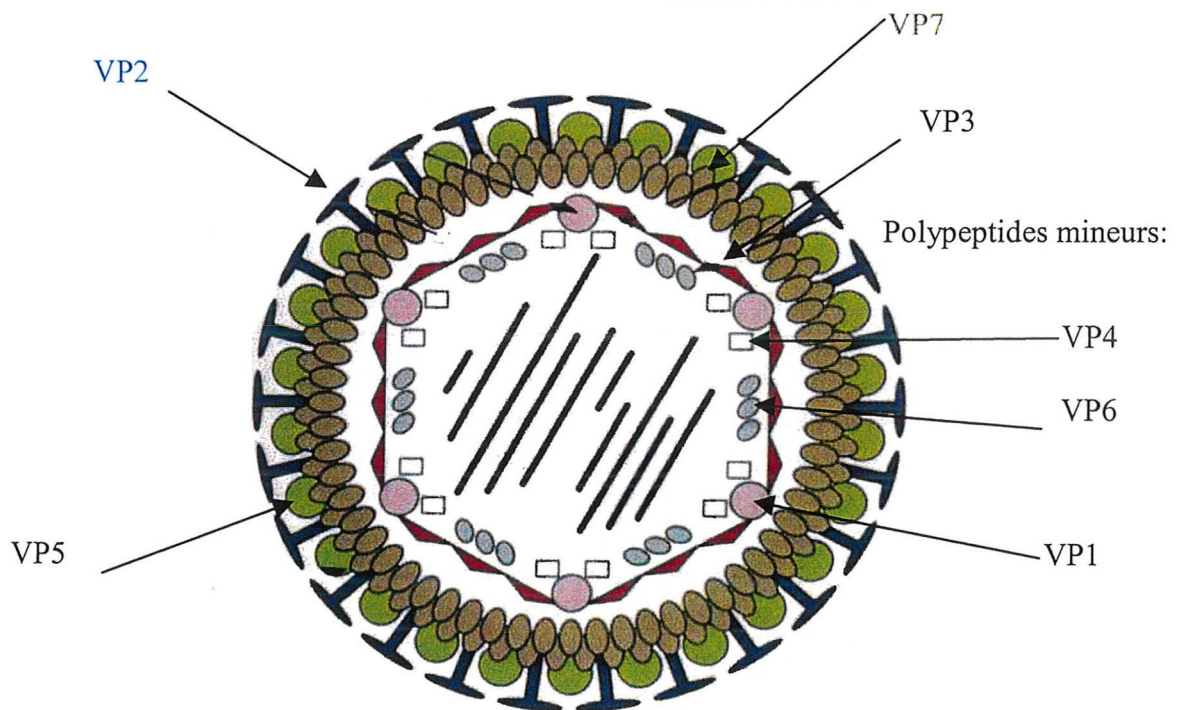
Membrane externe :

Nucléocapside:

Polypeptides majeurs:

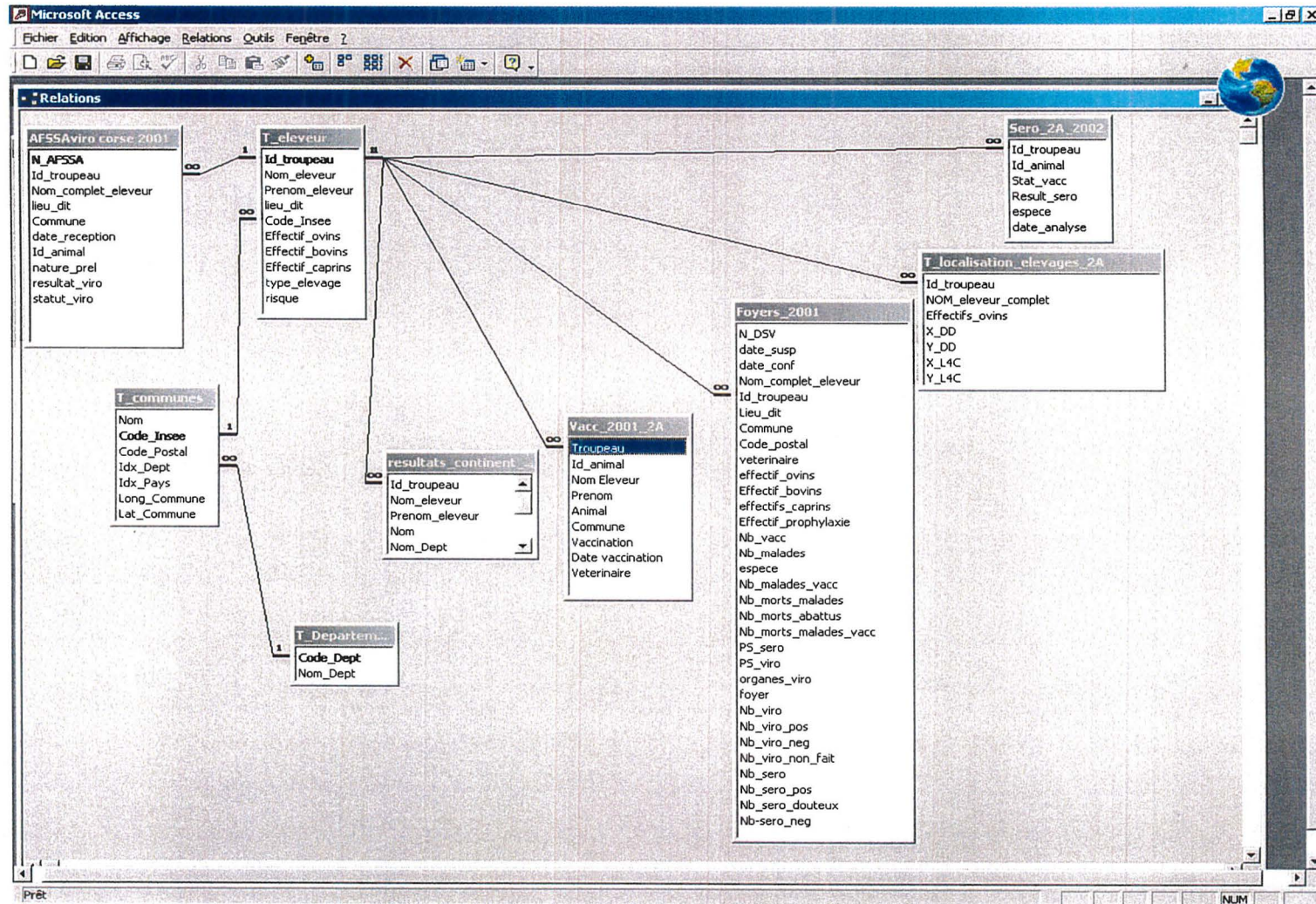
Anticorps spécifiques de type  
ex : SÉRONEUTRALISATION

Anticorps spécifiques de groupe  
Ex : cELISA



## ANNEXE IV

Base de données relationnelles créée pour la gestion des données sérologiques et virologiques collectées dans le cadre de la surveillance de la bluetongue en Corse





## ANNEXE V :

Détail des résultats de l'analyse ROC conduite en retenant dans l'échantillon les animaux non vaccinés dont le motif de prélèvement était la suspicion clinique

Cut-off	Sensibilité	Spécificité	correctement classifiés (%)
-68.06	100.00 %	0.00 %	61.63 %
-57.44	98.11 %	0.00 %	60.47 %
-38.29	98.11 %	3.03 %	61.63 %
-22.47	98.11 %	6.06 %	62.79 %
-20.86	96.23 %	6.06 %	61.63 %
-15.7	96.23 %	9.09 %	62.79 %
-14.99	96.23 %	12.12 %	63.95 %
-14.33	96.23 %	15.15 %	65.12 %
-13.59	96.23 %	18.18 %	66.28 %
-12.01	96.23 %	21.21 %	67.44 %
-11.38	96.23 %	24.24 %	68.60 %
-11.17	96.23 %	27.27 %	69.77 %
-10.78	96.23 %	30.30 %	70.93 %
-9.84	96.23 %	36.36 %	73.26 %
-7.43	94.34 %	36.36 %	72.09 %
-4.11	94.34 %	39.39 %	73.26 %
1.67	92.45 %	39.39 %	72.09 %
4.15	92.45 %	42.42 %	73.26 %
5.70	92.45 %	45.45 %	74.42 %
6.74	90.57 %	45.45 %	73.26 %
9.11	88.68 %	45.45 %	72.09 %
9.16	88.68 %	48.48 %	73.26 %
16.04	88.68 %	51.52 %	74.42 %
20.62	88.68 %	54.55 %	75.58 %
28.21	88.68 %	57.58 %	76.74 %
29.16	86.79 %	57.58 %	75.58 %
30.37	86.79 %	60.61 %	76.74 %
31.27	86.79 %	63.64 %	77.91 %
31.89	84.91 %	63.64 %	76.74 %
34.00	83.02 %	63.64 %	75.58 %
40.50	81.13 %	63.64 %	74.42 %
42.59	81.13 %	66.67 %	75.58 %
44.57	79.25 %	66.67 %	74.42 %
46.68	79.25 %	69.70 %	75.58 %
47.68	77.36 %	69.70 %	74.42 %
48.37	75.47 %	69.70 %	73.26 %
48.81	73.58 %	69.70 %	72.09 %
49.00	71.70 %	69.70 %	70.93 %
49.08	69.81 %	69.70 %	69.77 %
51.29	67.92 %	69.70 %	68.60 %
52.94	67.92 %	72.73 %	69.77 %
53.07	67.92 %	75.76 %	70.93 %
53.67	66.04 %	75.76 %	69.77 %
54.30	64.15 %	75.76 %	68.60 %
54.47	62.26 %	75.76 %	67.44 %
55.07	60.38 %	75.76 %	66.28 %
55.21	60.38 %	78.79 %	67.44 %
56.74	60.38 %	81.82 %	68.60 %
60.29	58.49 %	81.82 %	67.44 %

## ANNEXE V(suite) :

Cut-off	Sensibilité	Spécificité	correctement classifiés (%)
61.69	56.60 %	81.82 %	66.28 %
61.90	54.72 %	81.82 %	65.12 %
62.46	54.72 %	84.85 %	66.28 %
62.61	52.83 %	84.85 %	65.12 %
62.66	50.94 %	84.85 %	63.95 %
62.72	49.06 %	84.85 %	62.79 %
63.59	49.06 %	87.88 %	63.95 %
63.98	47.17 %	87.88 %	62.79 %
64.18	45.28 %	87.88 %	61.63 %
64.36	43.40 %	87.88 %	60.47 %
64.61	41.51 %	87.88 %	59.30 %
65.18	39.62 %	87.88 %	58.14 %
65.49	37.74 %	87.88 %	56.98 %
65.60	35.85 %	87.88 %	55.81 %
65.75	33.96 %	87.88 %	54.65 %
67.15	32.08 %	87.88 %	53.49 %
68.07	30.19 %	87.88 %	52.33 %
69.06	30.19 %	90.91 %	53.49 %
69.94	30.19 %	93.94 %	54.65 %
70.08	28.30 %	93.94 %	53.49 %
70.69	26.42 %	93.94 %	52.33 %
70.82	24.53 %	93.94 %	51.16 %
71.07	22.64 %	93.94 %	50.00 %
71.08	20.75 %	93.94 %	48.84 %
71.23	20.75 %	96.97 %	50.00 %
71.50	18.87 %	96.97 %	48.84 %
72.36	18.87 %	100.00 %	50.00 %
72.73	16.98 %	100.00 %	48.84 %
75.24	15.09 %	100.00 %	47.67 %
75.80	13.21 %	100.00 %	46.51 %
77.72	11.32 %	100.00 %	45.35 %
79.59	9.43 %	100.00 %	44.19 %
84.89	7.55 %	100.00 %	43.02 %
86.44	5.66 %	100.00 %	41.86 %
88.73	3.77 %	100.00 %	40.70 %
91.51	1.89 %	100.00 %	39.53 %
91.51	0.00 %	100.00 %	38.37 %



## ANNEXE VI :

### Manuel opératoire du test cELISA II



BLUE VMRD 001 - 110403

#### KIT VMRD BLUETONGUE, C-ELISA

**Kit de 5 plaques sécables (480 tests). Utilisable sur sérum ou plasma de ruminants.**

#### I - INFORMATIONS GENERALES

La maladie de la "langue bleue" est causée par un virus de la famille Reoviridae, genre Orbivirus. Elle est caractérisée par l'inflammation et la congestion pouvant entraîner la cyanose et l'ulcération des membranes muqueuses. Chez les moutons la Bluetongue se traduit par une forte morbidité et une faible mortalité.

#### II - PRINCIPE DU TEST

**Le Kit C-ELISA (Compétitive ELISA) VMRD Bluetongue détecte les anticorps anti-Bluetongue chez les ruminants.**

Les microcupules des plaques ELISA sont sensibilisées avec de l'antigène Bluetongue. Les anticorps spécifiques du virus de la Bluetongue et présents dans les sérums testés inhibent la fixation du conjugué anti Bluetongue marqué à la peroxydase. La fixation du Conjugué est révélée par un substrat enzymatique et quantifiée par l'apparition d'une coloration.

L'apparition d'une coloration indique la fixation du conjugué (absence de compétition avec le sérum) donc la présence d'un échantillon négatif en anticorps anti Bluetongue. **Une faible coloration, à l'inverse traduit la présence d'anticorps anti Bluetongue.**

#### III - COMPOSITION DU KIT

- **Réactif A : Plates coated with Bluetongue Antigen.** Microplaque ELISA sensibilisée avec antigène Bluetongue. 12 barrettes de 8 cupules.
- **Réactif B : Positive Control Bluetongue.** Sérum contrôle positif Bluetongue. 4,0 ml. Prêt à l'emploi.
- **Réactif C : Negative Control Bluetongue.** Sérum contrôle négatif Bluetongue. 4,0 ml. Prêt à l'emploi.
- **Réactif D : Antibody Peroxidase Conjugate.** Conjugué anti Bluetongue marqué à la peroxydase. Prêt à l'emploi.
- **Réactif E : 50x Wash Solution Concentrate.** Solution de lavage concentrée (50x). 60 ml.
- **Réactif F : Substrate Solution.** Solution substrat. Prêt à l'emploi. 30 ml.
- **Réactif G : Stop Solution.** Solution d'arrêt. Prêt à l'emploi. 30 ml.

#### IV - PRECAUTIONS D'EMPLOI

1. Lire attentivement la notice opératoire avant la réalisation des analyses.
2. **Conserver le kit et tous les réactifs entre +2 et +8°C.**
3. Porter les réactifs à température ambiante avant la réalisation des analyses.
4. Les barrettes inutilisées seront conservées entre +2°C et +8°C dans le sachet plastique fermé.
5. Ne pas mélanger des réactifs provenant de différents lots de réactifs.
6. Eviter la contamination des réactifs en utilisant du matériel de prélèvement à usage unique.
7. Ne pas utiliser de réactifs ayant dépassé leur date de péremption.
8. Ne pas manger, boire ou fumer durant la réalisation des analyses.
9. Utiliser un embout de pipette neuf pour chaque échantillon ou chaque contrôle.
10. Ne pas pipeter avec la bouche.
11. Utiliser les contrôles en double lors de chaque série d'analyses.
12. Utiliser uniquement de l'eau distillée (ou qualité supérieure) lors de la réalisation des analyses.
13. Eviter les contacts de la peau avec la solution substrat et la stop solution. Rincer abondamment à l'eau en cas de contact avec la peau.
14. Kit à usage uniquement vétérinaire.

## V - MATERIEL NECESSAIRE NON FOURNI DANS LE KIT

1. Micro pipettes de précision.
2. Embouts de pipettes à usage unique.
3. Eau distillée (ou qualité supérieure).
4. Micro pipettes multicanaux.
5. Centrifugeuse (2000\*g).
6. Tubes à centrifuger et microtubes.
7. Vortex ou équivalent.
8. Eprouvette d'un volume de 100 ml pour la solution de lavage.
9. Lecteur ELISA équipé d'un filtre à 620 nm.
10. Timer
11. plaques de transfert( plaques 96 puits fond en U).

## VI - INFORMATIONS SUR LES ECHANTILLONS

### Sérum ou plasma individuel : Les échantillons à analyser sont à déposer purs (sans dilution).

Des échantillons frais, réfrigérés ou congelés sont utilisables. Un volume de **25 µl** par échantillon est utilisé dans le test.

- Dans le cas où un nombre limité (une ou 2 barrettes) d'échantillons est testé, la distribution des échantillons et des contrôles sera réalisée directement dans la Plaque sensibilisée. Le Volume par échantillon ou par contrôle sera de **25 µl**.
- Dans le cas où un nombre important d'échantillons est testé, il est conseillé de préparer une « plaque transfert » (Plaque vierge). Distribuer **40 µl** de Contrôle Positif, de Contrôle Négatif et de sérum (ou plasma) à tester dans la « plaque transfert ». A l'aide d'une pipette multicannal de **25 µl** (8 ou 12 canaux), transférer les contrôles et les échantillons de la « plaque transfert » à la plaque sensibilisée.
- les sérums sont alicotés ( 200µl par puit) sur des plaques 96 puits, de même que les contrôles positifs et négatifs.

## VII - PREPARATION DES REACTIFS

- **Contrôles positif et négatif, Conjugué, Solution substrat, Solution d'arrêt :** prêts à l'emploi
- **Solution de lavage concentrée (50x) :** Solution à diluer au **1/50 dans de l'eau distillée** (ou de qualité supérieure) (**eau ultra-pure**)
- Exemple : 2 ml de **Solution de Lavage Concentrée (50x)** dans 98 ml d'eau distillée. Agiter après dilution. La solution de lavage diluée est à conserver au maximum 8 jours à +4°C.

## VIII - MODE OPERATOIRE

1. Porter les réactifs à température ambiante (18-25° C) avant la réalisation du test. Prendre une microplaque sensibilisée (Réactif A).
2. Les échantillons et les contrôles sont à analyser purs (sans dilution).
  - Déposer **25 µl de Positive Control Bluetongue pur** (Réactif B) dans les cupules A1 et A2 (par exemple).
  - Déposer **25 µl de Negative Control Bluetongue pur** (Réactif C) dans les cupules B1 et B2 (par exemple).
  - Déposer **25 µl de sérum (ou plasma) pur à analyser** dans les cupules suivantes.
  - Déposer 25 µl d'un sérum positif de référence dans les cupules E1 et E2 (pur), F1 et F2 (sérum dilué au 1/2) , G1 et G2 (sérum dilué au 1/4) et H1 et H2 ( sérum dilué au 1/8)
  - Les sérums à tester sont déposer en double ,cupules A3 et B3 pour le premier sérum et ainsi de suite.
3. Couvrir la plaque . Incuber la plaque 15 minutes à température ambiante (18 à 25°C) sous agitation douce.
4. Distribuer dans chaque cupule 25 µl de Conjugate Solution (Réactif D). Agiter pour bien mélanger
5. Incuber la plaque 15 minutes à température ambiante (18 à 25°C) sous agitation douce.
6. Vider la plaque et laver la plaque avec le laveur de plaques (3 lavages avec 200µl de solution de lavage). Vérifier l'état de propreté de la plaque et refaire un nouveau lavage si la plaque n'est pas absolument propre. Après les lavages, assécher la plaque en la tapant sur du papier absorbant.
7. Distribuer dans chaque cupule **50 µl de Substrate Solution** (Réactif F). Couvrir la plaque d'un papier d'aluminium.
8. Incuber **10 minutes** à température ambiante (18 à 25°C) sous agitation douce.
9. Distribuer dans chaque cupule **50 µl de Stop Solution** (Réactif G). Distribuer la **Stop Solution** dans le même ordre que celui de la **Substrate Solution**.
10. Lire la plaque à 620 nm au spectrophotomètre



## IX - CALCUL DES RESULTATS

Calculer la **DO moyenne (DOm CN)** du contrôle négatif.

Calculer la **DO moyenne (DOm CP)** du contrôle positif.

**Calculer le % Inhibition (% inh) des échantillons et du contrôle positif (CP) :**

$$\% inh \text{ Echantillon} = 100 - [(DO \text{ Echantillon}) / DOm \text{ CN}] * 100$$

## X - VALIDATION DU TEST

Le test est validé si :

$0,300 < DOm \text{ CN} < 2$
$\% Inh \text{ CP} \geq 50 \%$

## XI - INTERPRETATION DES RESULTATS

Données par le fabricant :

	INTERPRETATION
NEGATIF	$\% Inh \leq 50 \%$
POSITIF	$\% Inh > 50 \%$

Modification par le CIRAD-EMVT (introduction d'une zone douteuse)

	INTERPRETATION
NEGATIF	$\% Inh < 45$
DOUTEUX	$45 \leq \% Inh \leq 55$
POSITIF	$\% Inh > 55$